

Query Match	99.7%	Score 2024:	DB 6:	Length 2031:
Best Local Similarity	100.0%	Pred. No. 0:		
Matches 2031:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0
1	TTGGCAGCTGCAGAAACCAATCATCTCTCAGCTCCGCGCCGAGCCTTGAA	60		

Db 1 TTGCCATGCTGACGAAACAAATCAATCATCTCTGAGCTCCGCCACCCCTTGATATA 60
Qy 61 TGAATCTTAAGCAGACAGAGAGAAACATTCGGCAATCGCGAGCTGATGATACCTGAACA 120
Db 61 TGAATCTTAAGCAGACAGAGAGAAACATTCGGCAATCGCGAGCTGATGATACCTGAACA 120
Qy 121 CAAGCCTCATATGATTAAGATGAACAAACCTGGCCACAGTACTGAATGAGCCCTGGGGAA 180
Db 121 CAAGCCTCATATGATTAAGATGAACAAACCTGGCCACAGTACTGAATGAGCCCTGGGGAA 180
Qy 181 GGCCTTTCTATCCAAAGAAAGTATGTGGCAGCCGACACCTTTACAGTCACTATTAAAGAA 240
Db 181 GGCCTTTCTATCCAAAGAAAGTATGTGGCAGCCGACACCTTTACAGTCACTATTAAAGAA 240
Qy 241 GATGTCAAAAAGGCGTGTGGCACTGGAGTGAAGCCATTTCTCAATCACTCAGTTCAT 300
Db 241 GATGTCAAAAAGGCGTGTGGCACTGGAGTGAAGCCATTTCTCAATCACTCAGTTCAT 300
Qy 301 GACAAATATGATCAGATCCTTGAAGCCTGGAACGCATCGTGAACGCTGAGGACGCA 360
Db 301 GACAAATATGATCAGATCCTTGAAGCCTGGAACGCATCGTGAACGCTGAGGACGCA 360
Qy 361 CCTCTATCTCTGCAAGAGTTGAAGATCAAGAAACAGATCAGTGAATAAAGAAATGTG 420
Db 361 CCTCTATCTCTGCAAGAGTTGAAGATCAAGAAACAGATCAGTGAATAAAGAAATGTG 420
Qy 421 TCAGTGAACAATGGAAGGCTAGACGCGTGTATGAACTCTTAAACAGAGGAGAGAGAA 480
Db 421 TCAGTGAACAATGGAAGGCTAGACGCGTGTATGAACTCTTAAACAGAGGAGAGAGAA 480
Qy 481 ATGATGCTAGATCTGGGGGGACTGATTAAGACATATCTGSCAAAGCTGTTCAAGATAAG 540
Db 481 ATGATGCTAGATCTGGGGGGACTGATTAAGACATATCTGSCAAAGCTGTTCAAGATAAG 540
Qy 541 CTGGACCAAAATGCTTTTCAATTTGGAGAACATACACACACTGCTGGAAGAGAGAAAGCC 600
Db 541 CTGGACCAAAATGCTTTTCAATTTGGAGAACATACACACACTGCTGGAAGAGAGAAAGCC 600
Qy 601 AAACCTCTGAGATGATGAGAGCTAGACAGAAAGTTTCTGCTGATCATATGATGATA 660
Db 601 AAACCTCTGAGATGATGAGAGCTAGACAGAAAGTTTCTGCTGATCATATGATGATA 660
Qy 661 GTTACCATTTAAAGATATCTCAAGATTTCAATCCGGGACCTGGAAGATCTGGAATTCCT 720
Db 661 GTTACCATTTAAAGATATCTCAAGATTTCAATCCGGGACCTGGAAGATCTGGAATTCCT 720
Qy 721 TCAGTATGTAACCAACAGCAAGCAAGCAGAGACCATTAAGGGAAGAAATGATGAGACTA 780
Db 721 TCAGTATGTAACCAACAGCAAGCAAGCAGAGACCATTAAGGGAAGAAATGATGAGACTA 780
Qy 781 CAGGAGAGAGCTGATATGATTAACCTAGAGTTCTGAACCTCATTCGGGAGATGAGGAG 840
Db 781 CAGGAGAGAGCTGATATGATTAACCTAGAGTTCTGAACCTCATTCGGGAGATGAGGAG 840
Qy 841 CCTGATAAACCAATGCTCAAGAGATATCATAGATTAATTCAGAGATGAGATTCCT 900
Db 841 CCTGATAAACCAATGCTCAAGAGATATCATAGATTAATTCAGAGATGAGATTCCT 900
Qy 901 AAATAAAGCTTGAAGAACCGGATGACCAACTTGAAGCAATGACAGCTCCCTCAGTAC 960
Db 901 AAATAAAGCTTGAAGAACCGGATGACCAACTTGAAGCAATGACAGCTCCCTCAGTAC 960
Qy 961 AGATGAGCTGACAGGCGTATTTTCTGGGAGATATTCGACGAGTGAAGTTCGCTTCAT 1020
Db 961 AGATGAGCTGACAGGCGTATTTTCTGGGAGATATTCGACGAGTGAAGTTCGCTTCAT 1020
Qy 1021 GTCTCAATTTGAACAGCATCTCGAAATCTCAAGCAGCAGATTTGATGCTGAGACAAT 1080
Db 1021 GTCTCAATTTGAACAGCATCTCGAAATCTCGAAATCTCGAAATTTGATGCTGAGACAAT 1080
Qy 1081 TTAAGTCTGAGGCTTATCAACAGCAGATGAGTGAAGCACTGACTCATCCAGACGCG 1140
Db 1081 TTAAGTCTGAGGCTTATCAACAGCAGATGAGTGAAGCACTGACTCATCCAGACGCG 1140

Db 1081 TTAAGTCTGAGGCTTATCAACAGCAGATGAGTGAAGCACTGACTCATCCAGACGCG 1140
Qy 1141 TTTTGTAAAGAAAGTAAAGAGAGAGACAAACACTGTCTAGACCCCAAAATG 1200
Db 1141 TTTTGTAAAGAAAGTAAAGAGAGAGAGACAAACACTGTCTAGACCCCAAAATG 1200
Qy 1201 ANCTGATATTTGATATGCTAGCTGAGAGAGAGAAATCATCAACAGACATTAACCTGGA 1260
Db 1201 ANCTGATATTTGATATGCTAGCTGAGAGAGAGAAATCATCAACAGACATTAACCTGGA 1260
Qy 1261 GGGTGTCTATTAATGAGCTGAGTCCAACTGCTGATGAGCTCTGATGCTGA 1320
Db 1261 GGGTGTCTATTAATGAGCTGAGTCCAACTGCTGATGAGCTCTGATGCTGA 1320
Qy 1321 CACACGCGAGGCTTGAATGAGCAGAAACCTGTGAGAGAGACCTTAAGCCATTTG 1380
Db 1321 CACACGCGAGGCTTGAATGAGCAGAAACCTGTGAGAGAGACCTTAAGCCATTTG 1380
Qy 1381 AAATTGAACCTTGCACAGCATCATATGCTCAAAATGATGATTAAGCCATCAGTCCACAG 1440
Db 1381 AAATTGAACCTTGCACAGCATCATATGCTCAAAATGATGATTAAGCCATCAGTCCACAG 1440
Qy 1441 TGGAGCCGTTAATTAAGCAGAAATGATCTAATGATCAAGTGCAGAGAGAAAGCAA 1500
Db 1441 TGGAGCCGTTAATTAAGCAGAAATGATCTAATGATCAAGTGCAGAGAGAAAGCAA 1500
Qy 1501 GCAACCTTGAACCAAGCTAGAGGTTTAAATCAACGCTGGGCAAAATGTTTGGAAAAA 1560
Db 1501 GCAACCTTGAACCAAGCTAGAGGTTTAAATCAACGCTGGGCAAAATGTTTGGAAAAA 1560
Qy 1561 CAGAACAAAGAGAGCAGCTGATGATGCTGCTGCGCCAGGCAAAAGGTTCCATGCGG 1620
Db 1561 CAGAACAAAGAGAGCAGCTGATGATGCTGCTGCGCCAGGCAAAAGGTTCCATGCGG 1620
Qy 1621 AAATTGAGATTTGAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 AAATTGAGATTTGAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Qy 1681 CGCTGGAGGTTTACCGGAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1681 CGCTGGAGGTTTACCGGAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Qy 1741 CTGCTTTGAGACTAAAGAGAAACATTAAGAGTCTGATGAGAGAAAGGCCAGAGATG 1800
Db 1741 CTGCTTTGAGACTAAAGAGAAACATTAAGAGTCTGATGAGAGAAAGGCCAGAGATG 1800
Qy 1801 TTGCAAGATGCCCAAAATCTGCAAGACAAATATTTGACCAAGACATTAATTAACCTG 1860
Db 1801 TTGCAAGATGCCCAAAATCTGCAAGACAAATATTTGACCAAGACATTAATTAACCTG 1860
Qy 1861 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1920
Db 1861 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1920
Qy 1921 AATTCTAGACTGCGCGCGCGCGCTGAGAGCCCTTAATGATGCTGATTAAGATGAA 1980
Db 1921 AATTCTAGACTGCGCGCGCGCGCTGAGAGCCCTTAATGATGCTGATTAAGATGAA 1980
Qy 1981 TCACTAGTGGCGGCTGCAAGTCAATATGAGAGAGCTCCCAAGCGCT 2031
Db 1981 TCACTAGTGGCGGCTGCAAGTCAATATGAGAGAGCTCCCAAGCGCT 2031

RESULT 2
AB018271
LOCUS AB018271 5680 bp mRNA linear PRI 10-MAY-2002
DEFINITION Homo sapiens mRNA for KIAA0728 protein, partial cde.
ACCESSION AB018271
VERSION AB018271.2 GI:20521140
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCES
JOURNAL
COMMENT
FEATURES
SOURCE

1
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Nagase, T., Ishikawa, K., Suyama, M., Kikuno, R., Miyajima, N.,
Tanaka, A., Kotani, H., Nomura, N. and Ohara, O.
Prediction of the coding sequences of unidentified human genes. XI.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro
DNA Res. 5 (5), 277-286 (1998)
9872452
2 (bases 1 to 5680)
Ohara, O., Suyama, M., Nagase, T., Ishikawa, K. and Kikuno, R.
Direct Submission
Submitted (08-OCT-1998) Oasuu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology, Yana 1532-3, Kisarazu, Chiba
252-0612, Japan (E-mail: cdna@info.kazusa.or.jp, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)
On May 9, 2002 this sequence version replaced gi:3882176.
Location/Qualifiers
1. 5680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fn22477"
/sex="male"
/tissue_type="brain"
/clone_lib="pbluescriptII SK plus"
/dev_stage="adult"
/note="This sequence was replaced that of hK03530 cDNA as
a representative cDNA sequence for KIA0728."
1. 5680
/gene="KIA0728"
1. 4914
/gene="KIA0728"
/codon_start=1
/product="KIA0728 protein"
/protein_id="BAA34448.2"
/db_xref="gi:20521141"
/translation="EELMPWLTETOSIISOLPALAEYETLRQOEHRRLRELIASH
KPHIDKNKTPQPLELSPGEGFISIOEKVYADLTYSQIKEDVKRVALDEALISQST
QFHDKIDQIIESLESLIVERLQPSISAEVEKIQEISENQNVSVMDEKLPLETLEK
QGESEMLARSGGTGDKDISAKAVQDKLDQMTWENITLVREARVAKLDDVETLEKPM
CDHMSLIVTKIDTODPFIREDLEPDIDPSVVOQOEAETREIBEDYIOBELDIYNG
SELIACGEPDPKIVKSIDELINSLINKAKWDRIKLEEMQAAVQODGQANP
DMVDIAGKASNSPIGTDETVKQTEBELKQSEAYQOQIEERLNHQBELKLV
TEESDKTVODPLMELKILMSLEERIINRQKLEGALAGQPHALDELAVLTHT
EGLSEBQPVGGDKPALEIELAKHVLDQNDVLAQSTVEANAKGNDLIBSAGEBAS
NLQKLEVLNORQNVLEKTEQRKQDGLRQAKGPHGEIEDLQWLTDERHILAS
KPLGGLPETAKEOLNMEVCAAFEAKEERTYSIMQGOQWLARCPSAETNIDODIN
NIKEKMSVETKNERKTKEALNLAEMEPHNSIQDPIINMTQAEQTLNANSPSLIL
DTVLFQIDENKVPANESVSHREQIIEIDKTGTHLYFSQKQDVPLIKNLLISVGRNE
KVVQRLIVERGSLDDARKRAKQFHEAWSKLEWLESEKSDLELAINPDKIKTDL
AOHKEFOKSLGAKHSVYDTTRTGRSLKEKTSIADNDLKDDMLSELDKWDICGS
VERONKLEBALFSGQFDALQALIDMLRYVEPOLADQPVHGIDILVMNLDINHKA
KQELGKTSVQALKESARELIEGSRDSSVYQOMELSTRMTVLAISIKOTRLE
AALRQAEERHSVYVALLLEWLAEGQTLRPHGVLPDDEDAATLIDQKEPKKIEER
AELNKATITMGTVLAICHPSITTIKMTITIRAFEEVLAWAHQOORLASALAGLI
AKOELKRALLAWLQMAETTLTDKQKEVIPOEIEVKALIEHQTFMEEMRKQDVK
VTKTKRRALADPSSLOSHIPLYLDKGRGKRPASISYPSGQTOITKQIPRVVLVS
KMOOVLLALERRKLDNDALDRLEEFANFPDIWRKKYMRMNRKHSKSVDFPR
IDKQDQKITRROEFDIGLISKPSPTSLRLEMSAVADIDRDGDYIDYVEPVALHPMK
DAYKRTIDAKTIDBEVTRQVAKKCAKRFQVQJGDNKRYRFLDNGDSQQLVYAI
LRSTVNRVGGMRALDEFLVKNPCBARKGTNNELEKFLIADGASQGMARFPRR
RSRPSRSGASNRSTSVSSQAAQASPOVPATTPKILHPLTRYGKPMILNSMSTP
KKAECSDDFVPSAEGTPIQGSKLRLPGYLSGKGFHSGDGLTTAAARVITPAAS
KTPSRGSRAGSAGSRASRRSGSDAIDFISIGVCSDETVTPQTHRTPRAGSR
PSTAKPKIPIPPQKSPASKLDSKSR"

ORIGIN
Query Match 81.5%; Score 1655.2; DB 9; Length 5680;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1804; Conservative 0; Mismatches 73; Indels 18; Gaps 9;

Oy 8 GCTGACAGAAACAAATCAATCATCTCTCAGC--TCCGCCCCAGCCCTTGAATATGATAC 66
Db 18 GCTGACAGAAACAAATCAATCATCTCTCAGCTTCCGCCCCAGCCCTTGAATATGAAAC 77
Oy 67 TCTAAGCAGCAGCAGGAGAAACATCTGCACATCTGCTGAGTTGATGATCTGAACACAGCC 126
Db 78 TCTAAGCAGCAGCAGGAGAAACATCTGCACATCTGCTGAGTTGATGATCTGAACACAGCC 137
Oy 127 TCAATATGATATGATGAACAAACTGG----CCACAGTACTGAAATGAGCCCTGGGGAAAG 182
Db 138 ACATATATGATATGATGAACAAACTGGGGCAAGTTACTGAAATGAGCCCTGGGGAAAG 197
Oy 183 CTTTTCATCAAGAGAAAGATATGAGGAGCCGACACCCCTTACAGTACATTAAGAAGA 242
Db 198 CTTTTCATCAAGAGAAAGATATGAGGAGCCGACACCCCTTACAGTACATTAAGAAGA 257
Oy 243 TGTCAAAAGCGTCTGTGCACTGATGAAAGCCATTTCATCAATCAATCAATGATCAATGA 302
Db 258 TGTCAAAAGCGTCTGTGCACTGATGAAAGCCATTTCATCAATCAATCAATGATCAATGA 317
Oy 303 CAAGATGATCAATCTCTGAGAGCTGAAACCGATCTGTGAACCTGTGAGGACGCCACC 362
Db 318 CAAGATGATCAATCTCTGAGAGCTGAAACCGATCTGTGAACCTGTGAGGACGCCACC 377
Oy 363 CTCTATCTCTGAGAGGTTGGAAGATCAAGGAACAGATCAGTGAATAAATGATGTGTC 422
Db 378 CTCTATCTCTGAGAGGTTGGAAGATCAAGGAACAGATCAGTGAATAAATGATGTGTC 437
Oy 423 AGTAGACATGAAAGCTAAGCCGTGTATGAAACTCTTAAACAGAGGGAGAGAAAT 482
Db 438 AGTAGACATGAAAGCTAAGCCGTGTATGAAACTCTTAAACAGAGGGAGAGAAAT 497
Oy 483 GATTGCTAATCTGCGGGGGGACTGATTAAGACATATCTGCCAAAGCTTTCAGATTAAGCT 542
Db 498 GATTGCTAATCTGCGGGGGGACTGATTAAGACATATCTGCCAAAGCTTTCAGATTAAGCT 557
Oy 543 TGACCAATAGGTTTCAATTTGGGGAACATACACACCTGTTGGAAGAGAGGAGCA 602
Db 558 TGACCAATAGGTTTCAATTTGGGGAACATACACACCTGTTGGAAGAGAGGAGCA 617
Oy 603 ACTACTGAGTGTAGTGAAGCTAGCAGAAAGTTCTGCTGTGATCAATCATGATGATAGT 662
Db 618 ACTACTGAGTGTAGTGAAGCTAGCAGAAAGTTCTGCTGTGATCAATCATGATGATAGT 677
Oy 663 TACCATTAAGATATCTCAAGATTTCAATCCGGAGCTGGAAGATCTCGAAATGATCTCTTC 722
Db 678 TACCATTAAGATATCTCAAGATTTCAATCCGGAGCTGGAAGATCTCGAAATGATCTCTTC 737
Oy 723 AGTAGTAAACCAACAGCAAGAGAGAGACCATTAAGGAAATTAAGATGAGACTACA 782
Db 738 AGTAGTAAACCAACAGCAAGAGAGAGACCATTAAGGAAATTAAGATGAGACTACA 797
Oy 783 GAGAGAGCTGATATGATTTAACTGAGGTTCTGAAGCTATTCGCGGCACTGTGGGAGCC 842
Db 798 GAGAGAGCTGATATGATTTAACTGAGGTTCTGAAGCTATTCGCGGCACTGTGGGAGCC 857
Oy 843 TGATTAACCCATGTCGAAGAGATATACATGAGTTAAATTAACGAGTGGGATTTCTTAA 902
Db 858 TGATTAACCCATGTCGAAGAGATATACATGAGTTAAATTAACGAGTGGGATTTCTTAA 916
Oy 903 ATTAAGTTGGAAGACCGGATTTGAACCACTTGAAG-----AGCAATGACGCTCCGTCA 956
Db 917 ATTAAGTTGGAAGACCGGATTTGAACCACTTGAAG-----AGCAATGACGCTCCGTCA 976
Oy 957 GTACAGATGAGACTGAGAGCGGATTTTCTGAGTATATTTGAGTGTGATGATGATGATGAT 1016
Db 977 ACCAGATGAGACTGAGAGCGGATTTTGAATGAGTATTTGAGGATGATGATGATGATGAT 1036
Oy 1017 CAATGCTCAATTTGGAACAGATCTGGAACCTGTCAGAGAGAGATGATGATGATGATGATGAT 1076
Db 1037 CAATGCTCAATTTGGAACAG-ATCTGGAACCTGTCAGAGAGAGATGATGATGATGATGATGAT 1095
Oy 1077 CAATTTAAGTCTGAGGCTATCAACAGAGATGAG-AGTGTGAGAGTCACTCATCAGCA 1135

Db	1096	CAATTAAAGCTAGAGGCTATCAACAGCAGATGAATGAAAGACTGATCATCAAGCA	1155
Qy	1136	GACGCTTTTCTAAGAAAGTAACAGNAGAGTGAACAAACACACTGTTAGCCCAAA	1195
Db	1156	GA-GCTTTTCTTAAAGAAAGTAACAGNAGAGTGAACAAACACACTGTTAGCCCAAT	1214
Qy	1196	AATGAGCTGATATGTGATAT-GGCTAGCTTGAGNAGATCATCAACAGACATTA	1254
Db	1215	AATGAACTGAATATGATATGAGATAGCTTGAGAGAGATCATCAACAGACATTA	1274
Qy	1255	ACTGAGGGGCTCTATTAAGCCTGAGTCACTTCAACATGCCCTGTATAGCTCTG--	1312
Db	1275	ACTGAGGGGCTCTATTAAGCCTGAGTCACTTCAACATGCCCTGTATAGCTCTGCG	1334
Qy	1313	CATGTCGACACACAGGAGGCTTGCTAAGTGAAGCAAACTGTTGAGAGACCTTA	1372
Db	1335	ATGGCTGACACACAGGAGGCTTGCTAAGTGAAGCAAACTGTTGAGAGACCTTA	1394
Qy	1373	AGCCATTGAAATTTGAATCTTCCAGCATCATGTGCTTCAAAATGATGTATTAGCCCATCA	1432
Db	1395	AGCCATTGAAATTTGAATCTTCCAGCATCATGTGCTTCAAAATGATGTATTAGCCCATCA	1454
Qy	1433	GTCACAGTGAAGCCGTTAATTAAGCAGAAATGATCTAATTTGAATCAAGTCAAGAGAGA	1492
Db	1455	GTCACAGTGAAGCCGTTAATTAAGCAGAAATGATCTAATTTGAATCAAGTCAAGAGAGA	1514
Qy	1493	AGAAGCAGCACTTCAGAAAGAGCTAGAGGTTTAAATCAACGCTGGGCAAAATGTTT	1552
Db	1515	AGAAGCAGCACTTCAGAAAGAGCTAGAGGTTTAAATCAACGCTGGGCAAAATGTTT	1574
Qy	1553	GGAAGAAACAGAACAAAGAGAGCAGCTGATGTGCTTCCGCGCAGGCAAGAGGTT	1612
Db	1575	GGAAGAAACAGAACAAAGAGAGCAGCTGATGTGCTTCCGCGCAGGCAAGAGGTT	1634
Qy	1613	CGATGCGGAAATTTGAGATTTGACAGCATGTGCTGATGACAGCAGAGCTATCTTGTGCG	1672
Db	1635	CGATGCGGAAATTTGAGATTTGACAGCATGTGCTGATGACAGCAGAGCTATCTTGTGCG	1694
Qy	1673	ATCTTAAACGCTGGAGAGTTTACCGGAAACAGCAGCAGAGCAGCTTAAATGTCATTTGGA	1732
Db	1695	ATCTTAAACGCTGGAGAGTTTACCGGAAACAGCAGCAGAGCAGCTTAAATGTCATTTGGA	1754
Qy	1733	AGTCTGTGCTGCTTTGAAAGCTAAGAGAAATATTAAGTCTGATGACAGAAAGGCCA	1792
Db	1755	AGTCTGTGCTGCTTTGAAAGCTAAGAGAAATATTAAGTCTGATGACAGAAAGGCCA	1814
Qy	1793	GCGATGCTTTCGAAAGATGCCCCAAATCTGACAGACAAATATTGACCAAGACATTAATTA	1852
Db	1815	GCGATGCTTTCGAAAGATGCCCCAAATCTGACAGACAAATATTGACCAAGACATTAATTA	1874
Qy	1853	CTTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	1887
Db	1875	CTTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	1909

RESULT 3
AP400226
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

AP400226 16282 bp mRNA linear Pri 19-FEB-2002
Homo sapiens bullous pemphigoid antigen 1 ea (BPAG1) mRNA, complete cds.
AP400226 AP400226.1 GI:18157650
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Okumura, M., Yamakawa, H., Ohara, O. and Owaribe, K.
Novel alternative splicings of BPag1 (bullous pemphigoid antigen 1) including the domain structure closely related to MacP (microtubule actin cross-linking factor)

JOURNAL J. Biol. Chem. 277 (8), 6682-6687 (2002)
MEDLINE 21839111
PUBMED 11751855
REFERENCE 2 (bases 1 to 16282)
AUTHORS Okumura, M., Yamakawa, H. and Owaribe, K.
TITLES Direct Submission
JOURNAL Submitted (18-JUL-2001) Grad. Human Informatics, Nagoya University, Unit of Biosystems, Nagoya 464-8601, Japan
FEATURES
source
1. 16282
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="6"
/cell_type="keratinocyte"
1. 16282
/gene="BPAG1"
1. 15516
/note="BPAG1a; MacP-like product; alternatively spliced"
/product="bullous pemphigoid antigen 1 ea"
/protein_id="AA162061.1"
/db_xref="GI:18157651"
/translation="MSSSYSSRSSDSVFSNTTSTRSLSDSNENLLVHCPTLINSCLIFGSGSPGHRLEMLQIYANRQVDSYCEKLLIAGNALQSDSRLESQVQFQWEAELIYECENLROHVIDVOILIDQKQVQOQLVORVAKLDEIMALENCSYVKGILITTEORTKMLISGTLQSLNSGTPVOTLPSLISGLTSLTSLSSMTSGLSGMSRLTSTVPTATTPGPGSLVFPSSGVTENSLQTLKQNRPLKSLSLDQNTTEEINMKTVDLLNVDENQVLDRTWEGSDLPVSHEENKKNVRAIIEBESSLKEAKTISEIOMTAPLKLTYARKHLRESQYAKLNTSVNORHLDLTNPFASRATNELIWLKEEBEEVAADWSEENNTIARKKDYHAKLMELEDKENRISVQEIAGLELLENHARLTI BAYRAMOTOMSMILOTCCQVEOHIKENTAYEPFNDAEADYLDLDAIORKSC DRSSEIHLKLEDDVOESMEKEBELLYKSTIATLMGAKTILQIKRNSCPKLTSPICAIQVROYETITTYDDECVLANNHRAKVIISPGNAMPVSCPTTPPNKEAYD LAMRLEQOQVNTLTHBSHINMKS VSNHVDLNEIDRLASNVASIKTMLBEEHQV LNSLRQREFDLEDSQESQVFSQSDITOLEKVNVCQYQYQYBLKSAEKEBEEQVY LYSERVNRRLLENCEBRLRIQITRPLERDILHNSVFRITPEOKKKEBELKQDLOG TTNKCEEFSSQAAASSVPLRLSELNVLLQNMNVNMSSTYIDKLKTVNLKNTQ AABALVKLYETKCEBEAVIADKNINBNLISLTKOMRSEVDKROFHALDELOKAT AISDEMTYKERRDLPFMHKEKAQVLYEROMYVQIDNRJRDLEIGIKSLKYRPT YHPLDWMIOQVETTRKTIQENPENSKITATQUNQKMLVSEIEMKQSFNDICQKTA E QYATVEDELQITMTYRAMVDSQKSPVRRKMQSSADLIQEFMDLRRTYALVTLM TOYIKFAGDSLRLLEBEEKSLBEEKKEHVEKKEKELQKWSNISKTLQDAEKAGKPPS KOKISSEESTKQDLEALOTIOLFLAGKGMPTDEBENLEKQVTKLOESYNLFS BSLKLOESQTSQDVYKVEKTYAERQOEKEXLQGIQDILQOTENRLLGHQAPMIGD GTVELKTKOSKQEBELQKMQSSAQALAEVYKATNPFLENKGLSIOEDBALIYQKLN E AKIKCEQMLNKAQSKKELEIDVNTTAIKETKVAVAQVLEBSKTKIENLMLNSVD KOSERAGTHKQVIEQNGTHFQEGGKSAIGEDEVNGLRLTDVQGVGTQENLNO QYQKYVAQHEKTIISOHQA VILATQSAQVLELQOGQYLSPBEEKELQKMKLEKHYET ALAEBEKMKLTHSLOELEKFPDADYTEBHMLOQSEOLENLLEAGADINGLMFKLK ROKRSFSDVISHKGLRYITTSNGRVLEAKSCSRDGGKQVTSATHPVOKLQAHAT DPFRSLKYKCVNLGNMLKDUDVYKQHYEDASCGLLAGDQACATSKSLSEIYADPK NLQRLQLETKALQEQIISQOQVAVELKTAEVLILARQSLAPARKDIOCTLTDIADRY EBDLSKSVNERNEKLOITLRLSLVDQDLEMENGNVSSLKEGOVPLNLTALQDI I SKSNMLEDDIAGROSSINAMERKYMETTPDSTASLQKMDLSARFSEASHIK KETTLAKMEKLVKLEFENLSEKLOTPLETKQVALDEVDPKQVTELSQYQOESGSE PLEHKHLELVLSILKETISSHGLPSDKLIVLEKTNLSKKPEMDTJKEKKEATSC QEQDLAFQVLYVKSLSKMIKETTKYPIYQPSFGAIDGKSLDLYTKLOKBSKLTPEI QKVNNSGISLNLISAVTTPAALIAVNSGGAIVLAGEATNTEEFMANKGTLTSIKD MTDISHGYEDLGLKDKIABLNTLSLQKQAOEBSASAMQLOKMNKTLTMOQTPA PRTDAVTVQVQONKSFEABELKQNVKQKQVLELQKTELEBENPDPPEAPRMQMLTEI DSKMOELNOLITDROOKLEESNNLTQFQVYALQKQVLEKELNVSVGLPSSIDPM LINTBOQOVLLOEFATRRKPOYEOQLTLAGGQILSPRGIVKQDLAIVTQMD SITGQLSDBRCMTDAIYKSTQYQSLASLSLSDLDLNRKSSSLAVSTHPRANQOL ETLAQKROBEIQQERKQIVQAQVLCEDLSALVEEYLVKELSLQSLGILSIFQVEOKA ENHVLQASACASHQFOQMSRDPQAMVLPKKEBQONKSPIDKADYVLESLLKDNHDF SKTLTLAGSHMYEKTLTAEGBNLTKQSEKALQOQLNTIKNMPTFPNVOVERENKL KSELKALKYKQVETLWPMIDCONNLEIIFCQDPAEGENSIKTLKSLQEMODHF GAVELIINTANSILSVCEIDKGVYTPDKSLIQXYDMPTEQOLSHSKFCELENNQKKE FQEVKESRQLOQCKEODLHDSLGQVYSKTYIPLMDTOOKSLOALHGYDLARL AODLVAESDSKSTGVDLVQVETLQAESHTLSQVDEKCSFETKLQGIQGHRRNTIRE MFSQFAEPDEIDMSMAVGRDAETLQOKETIKAFKLKLEALMASNDNANKTCQMLA

TEESTPDLVGIKRDLBALSKCNKLLDBAORBEVGEITIKLEBEFYSKLEFSILLO
 PAEBHEESQDVGWMEETETINOGLMFPVFOKEBIEPTQGRQODLVNMLTAEILSMTVEIL
 TSTQGLDEDDDVNAKRTINKKYAQAQOQEAALLHOCGRQODLVNMLTAEILSMTVEIL
 VANKPPSAEFKVVAKQIOEQKLOQLDLDBKSTVEYIKREGEKIATTAEBLADPKVLL
 KQSLDLRMEALNKAETRNROLEGISVAVQOQHELETLEPENMTTIEKLVNCEBI
 QTOASKEBQIAQKLEDDIINHNKILQAVSIGOSIKLVNSBEDDMQVSKLDFSQ
 VVYIEIOEKSHSRSEELLQALCNKIKFGEDEVELMMNVENYHDLSTADYDSTEGE
 WKOOSLRYLQEDILIRKONVDOALLNGLELTKOTDEVTILIDDKLEIKYARQDIT
 KLSIDVAKTLEQALQARLRLSHBELCTMLDKYVELLSTETVLGEBEASQOMRP
 KELEKKAENKNNKALDLSNEVSALLELPMPARBLEMVAENDEKRLVSDITTOYV
 BEVDAAILRSQODPOAADLMSWITTEKELMSGLDILBEDQMSAQLOVQKPTMEI
 LRHDIIDDLVYSGHKIMTACSEEEKOSMKKLDKVLNKYITICQINSERLAQERAO
 SLVQFWEYTEELMPLTETOSIISOLPAPLRYETIRQOEHKROEERLSEIAHKPHI
 DKWKTQPOLLELSRGEFSIOEKYVADUTYSOIKEDYKRAVALDBALISQOFPHD
 KIIDILSLEHIVRLKOPESISAEVKIKQISENKVSDMKLPLLETLKORSE
 EMILARSGETDIDISAKAVODLQDMVEITWENIHVLVEERKMLDVELAEKFCDH
 SLVITIKOTDPIIDLEGPIDPSVVKQOQAETIIBEIDGLEBEIDIVNLGSEL
 AACGEPKPIYKKSIDELNSAMDNLKAMKDRIDLEBAMQAAVOYODGLOAVFDWD
 IAGKLSAMSPIGTDLFTVKKOIBELKQFSEAVOOOIMEMERLQABELLKKTYES
 DKHTVODPMEIKLIWDSLEBRIINROHKFEGALLAQQOPHALDELLAMTTHREGIL
 SEQKPVGDPRALIEIELAKHVLQNDVLHOSITVEAVNKGNDLIESAGEASNLQ
 KLEVLNORWQNVLEVEVCAQKQOLDGALROAQFGEILEDLOQMLDTERHLLAKPLG
 GLPEITAKEOLNVMEVCAFAKEEYKSLQKQCOMLARCPSKAETNIIDQINLKE
 KWESVETKLENERKTLBEALNLAMEFNHSLDDFIMLTOAOTNLVARSPLIDTVL
 FOIDEHVPANEVNHSRBOILELKTGTHLKYSOKODVULIKULISVOSRMEKVVIO
 RLYBERGSLDARAKAQFHEANSKLMEMLEBSKSLDSELEIANDPKITTOAQK
 EFQSLAKHSVYDTNTRKSLNEKTSVLADNLLKLDMSLEKDKDITCGKSEVO
 NKLBEALLFSQGFDTALQALIDMLYRVEPOLABEQPVHGDIDLVNMLIDHNKAQKEL
 QKERTSVYOLAKRSABELIXGRDSSWVKVMOELSTRMEVTCALSTISKROKLEALR
 OAEFHSVYHALLMELEABEOTLRFHGLPDEBDALRLIDONHEPMKKEEKAEN
 KATMGDTVLAICHPDSTITIKHWITIIIRAFREVLAMAKOORLALAGLAKOE
 LLEBALLMLOWAETTLTLDKKEVYPOEIEBKALIAEQOTFMEKTKODPVDTYTT
 YKRPADPSSLOSHI PVLGDKRGRKRP PASLYPSGQTOIETKNPRLVLSKWOQ
 VMLALERRRLDNLDRLELREFANVEDIIMRKXMRWNHNSKMPFNRIIDAK
 QDGIKTOEPIIDILSSKFTPSRLSESAVADIPIRDBGYIDYVEFPAALHLLAKP
 PITVADGKIDEDVTROVAKCKCAKRFQVOIGDNKRYFLGQFDSOOLRLYLRST
 VMVRGGMMALDEFLVKNDBCRAGKRTNMLREKTLIADASOGMAAPRRGRSRSP
 SSRGASPRRSTSVSQAQAASPOVPATTTKILHPLTRYKRWLNSKSTCKKA
 EBSDPVPASAGTPIQGSKRLPGYLSGKHSEGBDGLITTAARVYTOPADSKTP
 SRPSRAGSKAGSRASRSRGSADDFISEIOWSVCSDEVTVPTRHPRAGSRPSRA
 KPSKIPPORKSPASKLDKSSKR"

ORIGIN

Query Match 81.3%; Score 1652; DB 9; Length 16282;
 Best Local Similarity 95.1%; Pred. No. 0;
 Matches 1802; Conservative 0; Mismatches 75; Indels 18; Gaps 9;

8 GGTGACAGAAACACAATCATCTCTCAGC-TCCGCGCCGAGCCCTGTAATATGATAC 66
 |||||
 Db 10620 GCTACAGAAACACAATCATCTCTCAGCTTCCGCGCCGAGCCCTTGAATATGAAAC 10679

67 TCTAAGCAGCAGCAGAGAAACATCCGCACTCGTAGTGTGATAGCTGAACACAGCC 126
 |||||
 Db 10680 TCTAAGGAGCAGCAGAGAAACATCCGCACTCGTAGTGTGATAGCTGAACACAGCC 10739

127 TCAATATGATATGAAACAAACTGCG---CCAAGTACTGAAATGAGACCCCTGGGGAGG 182
 |||||
 Db 10740 TCATATATGATATGAAACAAACTGCGCCACAGTACTGGAATTTGAGACCCCTGGGGAGG 10719

183 CTTTTCATCAAGAGAATATGAGCAGCAGCAGCAGCCCTTTCACGTCAATTAAAGAA 242
 |||||
 Db 10800 CTTTTCATCAAGAGAATATGAGCAGCAGCAGCAGCCCTTTCACGTCAATTAAAGAA 10859

243 TGTCAAAAAGCGTCTGTGAGCAGTGAAGCAGCATTTCATATCAATCACTCAGTTCAATGA 302
 |||||
 Db 10860 TGTCAAAAAGCGTCTGTGAGCAGTGAAGCAGCATTTCATATCAATCACTCAGTTCAATGA 10919

303 CAAATATGATCAGATCTTGAAGCCTTGAACGCATCTGTGAACGTTCTGAGGACGCCACC 362
 |||||
 Db 10920 CAAATATGATCAGATCTTGTGAGGCTTGAACGCATCTGTGAACGTTCTGAGGACGCCACC 10979

363 CTCTATCTGTGAGGTTGAGAGATCAAGAAACAGATCAAGTAAATAAGATGTGTC 422
 |||||

Db 10980 CTCTATCTGTGAGGTTGAGAGATCAAGAAACAGATCAAGTAAATAAGATGTGTC 11039

423 AGTGAACATGGAAGAAAGCTACACGCCGTTGTATGAACCTTTAAACAGAGGGGAGGAAAT 482
 |||||
 Db 11040 AGTGAACATGGAAGAAAGCTACACGCCGTTGTATGAACCTTTAAACAGAGGGGAGGAAAT 11099

483 GATTGACAGATCTGTGGGGGACGTGATTAAGACATATCTGCCAAACCTGTTCAGATTAAGCT 542
 |||||
 Db 11100 GATTGACAGATCTGTGGGGGACGTGATTAAGACATATCTGCCAAACCTGTTCAGATTAAGCT 11159

543 TGAACCAATAGGTTTCTCATTTGGAGAAACATACACACTGGTGAAGAGAGGAGCCAA 602
 |||||
 Db 11160 TGAACCAATAGGTTTCTCATTTGGAGAAACATACACACTGGTGAAGAGAGGAGCCAA 11219

603 ACTATGAGATGTGAGAGCTAGCAGAAAGTTCTGTGTGATCAATGTTCATTTGATAGT 662
 |||||
 Db 11220 ACTATGAGATGTGAGAGCTAGCAGAAAGTTCTGTGTGATCAATGTTCATTTGATAGT 11279

663 TACCATTAAGATCTCAAGATTTTCATCCGGGACCTGGAAAGATCTGGAATTGATCTTC 722
 |||||
 Db 11280 TACCATTAAGATCTCAAGATTTTCATCCGGGACCTGGAAAGATCTGGAATTGATCTTC 11339

723 AGTATGAACCAACAGCAGAGAGCAGAGACCATTAAGGGAAGAAATAGATGACTACA 782
 |||||
 Db 11340 AGTATGAACCAACAGCAGAGAGCAGAGACCATTAAGGGAAGAAATAGATGACTACA 11399

783 GAGAGACCTGAGATATAGTATTAACTTAACTTGAAGTTCTGAACTCATTCGCGCATGTGGGAGCC 842
 |||||
 Db 11400 GAGAGACCTGAGATATAGTATTAACTTAACTTGAAGTTCTGAACTCATTCGCGCATGTGGGAGCC 11459

843 TGAATAAACCAATGTCAGAAAGAGTATACATGAGTTTAAATTCAGCGAATGGGATCTCTAA 902
 |||||
 Db 11460 TGAATAAACCAATGTCAGAAAGAGTATGATGATTAATTAATTCACG-ATGGGATTCCTTAA 11518

903 ATAAAGCTTGAAGAAAGACCGGATTCGACCACTTGAG-----AGCAATGACAGCTCCGCTCA 956
 |||||
 Db 11519 ATAAAGCTTGAAGAAAGACCGGATTCGACCACTTGAGAGGCAATTCAGAGCTCCGCTTCAAGT 11578

957 GTACAGATGAGCTGACAGCGCGTATTTTCTGGGTAGATATTTGACAGTGTGATAGTTCCCTT 1016
 |||||
 Db 11579 ACCAGATGAGCTGACAGCGCGTATTTTCTGGGTAGATATTTGACAGTGTGATATTTAATTAAGCTT 11638

1017 CAATGTCCTCAATTTGGAACACAGATCTCGAAACCTGCAACGACAGATTTGATGTCTAGAG 1076
 |||||
 Db 11639 CAATGTCCTCAATTTGGAACAG-ATCTGGAACCTGTCAGAGCAGATTTGAAGACTTAAG 11697

1077 CAATTTAAGTCTGAGGCTTATCAACAGCAGATAG-AGTGTAGCAGTGACTCATCAGCA 1135
 |||||
 Db 11698 CAATTTAAGTCTGAGGCTTATCAACAGCAGATAGAGTGAATGGAAGACTGAATCATCAGCA 11757

1136 GACGCTTTTCTTAAAGAAAGTAAACAGNAGAGTGAACAAACACACTGTTCTAGACCCAAA 1195
 |||||
 Db 11758 GA-GCTTTGCTTAAAGAAAGTAAACAGNAGAGTGAACAAACACACTGTTCAAGCCCAATT 11816

1196 AATGAGNCTGATATTGATAT-GGCTAGCTGAGNAGAGATCATCAACAGACAGCTTAA 1254
 |||||
 Db 11817 AATGGAACGTAAATTTGATATGAGATAGCTGAGAGAGATCATCAACAGACAGCTTAA 11876

1255 ACTGAGAGGAGTCTTATTAAGCMTAGGCTCAGTTCCAAATGSCCTGTATGAGCTCCGT-- 1312
 |||||
 Db 11877 ACTGAGAGGAGTCTTATTAAGCMTAGGCTCAGTTCCAAATGSCCTGTATGAGCTCCGTGCGC 11936

1313 CATGCTGACACACACGAGGCTTGTCTTAAGTGAACAGAAACCTGTGTGAAGAGACCTTAA 1372
 |||||
 Db 11937 ATGCTGACACACACGAGGCTTGTCTTAAGTGAACAGAAACCTGTGTGAAGAGACCTTAA 11996

1373 AGCCATTGAATTTGAATCTTGCAAGACATCATGTGCTCCAAATGATATTAAGCCATCA 1432
 |||||
 Db 11997 AGCCATTGAATTTGAATCTTGCAAGACATCATGTGCTCCAAATGATATTAAGCCATCA 12056

1433 GTCCACAGTGAAGCGTTTATATTAAGCAGGAATGATCTAATTTGAATCAAGTGCAGGAGA 1492
 |||||
 Db 12057 GTCCACAGTGAAGCGTTTATATTAAGCAGGAATGATCTAATTTGAATCAAGTGCAGGAGA 12116

Db 2714 ACTGAGGAGGCTCTATTAGGCTTGGCTCAGTTCACATCGCTGATGAGTCTCTGGC 2773
Oy 1313 CATGCTGACACACCCGAGGCTTCTTAAGAGAGAGAAACCTGTGGAGAGAGCCCTAA 1372
Db 2774 ATGCTGACACACACCCGAGGCTTCTTAAGAGAGAGAAACCTGTGGAGAGAGCCCTAA 2833
Oy 1373 AGCCATTGAATTTGAATTTGGCAAGCATCATGTGCTCCAAATATGATGATTTAGCCCATCA 1432
Db 2834 AGCCATTGAATTTGAATTTGGCAAGCATCATGTGCTCCAAATATGATGATTTAGCCCATCA 2893
Oy 1433 GTCACAGTGGAGAGCCGTTAATTAAGCAGAAATGATCTAATTAATCAAGTGCAGAGAGA 1492
Db 2894 GTCACAGTGGAGAGCCGTTAATTAAGCAGAAATGATCTAATTAATCAAGTGCAGAGAGA 2953
Oy 1493 AGAAGCAGCAACCTTTCAGAACCACTTAAGAGTTTAAATCAAGCTTGGCAAAATGTTTT 1552
Db 2954 AGAAGCAGCAACCTTTCAGAACCACTTAAGAGTTTAAATCAAGCTTGGCAAAATGTTTT 3013
Oy 1553 GGAAGAAAAACAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1612
Db 3014 GGAAGAAAAACAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3073
Oy 1613 CCATGCGGAAATTGAGAGATTTCAGACAGTGGCTGACTGACACGAGAGCTCATCTTGGCC 1672
Db 3074 CCATGCGGAAATTGAGAGATTTCAGACAGTGGCTGACTGACACGAGAGCTCATCTTGGCC 3133
Oy 1673 ATCTAAACCCCTGGAGAGTTTACCGGAAACGCCAAGAGAGAGAGAGAGAGAGAGAGAG 1732
Db 3134 ATCTAAACCCCTGGAGAGTTTACCGGAAACGCCAAGAGAGAGAGAGAGAGAGAGAGAG 3193
Oy 1733 AGTGTGTGCTGCTTTGAGAGTAAAGAGAAACATATATAGAGTCTGATGACAGAAAGGCCA 1792
Db 3194 AGTGTGTGCTGCTTTGAGAGTAAAGAGAAACATATATAGAGTCTGATGACAGAAAGGCCA 3253
Oy 1793 GCAGATCTTGCAGAGATGCCCAAAATCTGCAGAGACCAATATTCAGCAAGACAT 1846
Db 3254 GCAGATCTTGCAGAGATGCCCAAAATCTGCAGAGACCAATATTCAGCAAGACAT 3307

RESULT 5 3307 bp mRNA linear PRI 30-JAN-2004
AK096713
LOCUS
DEFINITION Homo sapiens cDNA FLJ39394 fls, clone PLACE6007925, moderately
similar to Homo sapiens mRNA for actin binding protein ABP620.
AK096713
VERSION AK096713.1 GI:21756267
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Ohyaishi, M., Nishi, T., Shibahara, T., Tanaka, H.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosioki, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuno, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Oho, Y., Takiguchi, S., Watanabe, S.,
Yoshida, M., Hottel, T., Kusano, J., Kanehori, K., Takahashi, H.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Mutsaers, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunaga, H., Ichihara, T.,
Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Sato, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Teshiro, H., Tanigami, A., Fujimura, T., Ono, T., Yamada, K., Fujii, Y.,

Ozaki, K., Hira, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, K.,
Noguchi, S., Itoh, T., Shigeta, K., Senda, T., Matsunaga, T.,
Nakajima, Y., Mizuno, T., Norinaga, M., Sasaki, M., Togaishi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikkawa, H., Masuo, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T., and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikkawa, H., Murakawa, K., Kanehori, K., Takahashi, Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahari, K., Masuo, Y., Nagai, K., and Isogai, T.
NEBO human cDNA sequencing project
Unpublished
3 (bases 1 to 3307)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazuo-Kamata, K., Katsura, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5' - & 3' - end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

COMMENT

FEATURES
SOURCE

1. 3307
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE6007925"
/tissue_type="placenta"
/clone_lib="PLACE6"
/note="Cloning vector: pMB18SFL3"
336..3307
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC04848.1"
/db_xref="GI:21756268"
/translation="MWQSKDFFQWVYIEIOEKSHRSSELLQALCNKIFGEDEVET
MMNLEVHDLSTSLVODSTEGMLQSESLRLOESILLRKNVQDALNGLELQ
TTDDDEVLTIDKLEAIKARYKDTSTYDAKLTDLQALQALRSHLSHEELCTMDLV
EGLVLETFDYLVKGBEASQAKMPKELKSAKNNKALDLDSVSSALLELVMPRAB
GKEMVAEDNERLRLVSDITQVVEIDAILLSQDFQDADELDSWITETKLNLSL
GDLRLDQVQADLOVQKTFMTBLIRHDKIIDLVVSGHGMTRACSEERQSPKALRY
KVLNPDTCQINSEERYQLERASLVNQCWYIEELMPLTETQSTISQLPAPLBY
ETDROOEHRLOREILAEHPHIDKNNKGPOLLESPGEGSIOEKQVAAVPTLVQ
IKEDVCRVALDEALISQHPHIDKNNKGPOLLESPGEGSIOEKQVAAVPTLVQ
ENQVSDYNEKLOPLVETLKQREEMTASGGTQKDISAKAVDDKQDFNTENIT
LVBERKLDVDELVELEKFCWCDMSLIVTQKQDTRIDEDGIDPSVVKQOQEAAB
TIBERIDQOEELDIVINSEELIACGEDPKIVKKSIDELNSADSLNKAWKDRID
KLEBAQAQVQDQAVDWDVADIGAKSASPIGTLETQKQIIEELKQKSEAY
QOIEEMERLNAHQEELKKVTESSDGHVQDPLMEELKLWDSLEERIINROHLEGL
LALGQONHLDIELLAVLTTEGLISROKPGQDPAIELELAGHYLQNDVLAHQSTV
EAYRKGNSLIESASGSEASNLQNKLEVLNQRQNYLERTQKQDQDLGKRAKGRH
GRLIEDIQWTLTDERHLASKPLGIG.PETAKQELNVHMECAAFEAKEETYSKLMQKQ
QOMLACPKSAETNIDOD"

CDS

ORIGIN

Query Match

80.64; Score 1637.6; DB 9; Length 3307;

Best Local Similarity 95.9%; Pred. No. 0;
Matches 1778; Conservative 0; Mismatches 58; Indels 18; Gaps 9;

```

QY      8 GGTGACGAAACCAATCAATCATCTCTCAGC--TCCGCGCCCAAGCCCTTGAATATGATAC 66
Db     1457 GCTGACGAAACCAATCAATCATCTCTCAGCTTCCGCGCCCAAGCCCTTGAATATGAAAC 1516
QY      67 TCTAAGCAGCAGCAGGAAAGAACATCGGCACTCGTGAAGTTGATAGCTGAAACAAGACC 126
Db     1517 TCTAAGCAGCAGCAGGAAAGAACATCGGCACTCGGTAAGTTGATAGCTGAAACAAGACC 1576
QY     127 TCAATATAGATTAAGTGAACAAAACCTG---CCAAGTACTGAATGAGCCCTGGGAAAG 182
Db     1577 TCAATATAGATTAAGTGAACAAAACCTGGGCCACAGTTAAGTGAATGGCCCTGGGAAAG 1636
QY     183 CTTTTCATCTCAAGAGAGATGATGGGAGCCGACACCCCTTACAGTACACTTAAGAGAG 242
Db     1637 CTTTTCATCTCAAGAGAGATGATGGGAGCCGACACCCCTTACAGTACACTTAAGAGAG 1696
QY     243 TGTCAAAAAGCGTGTCTGTGCACTGATGAAGCCATTTCTCAATCAACTCAGTTCCATGA 302
Db     1697 TGTCAAAAAGCGTGTCTGTGCACTGATGAAGCCATTTCTCAATCAACTCAGTTCCATGA 1756
QY     303 CAAGATATGATCAATCTCTTGAAGCTGGAACGATCTGTGAAACGTCTGAGGACGCCACC 362
Db     1757 CAAGATATGATCAATCTCTTGAAGCTGGAACGATCTGTGAAACGTCTGAGGACGCCACC 1816
QY     363 CTCATCTCTGCAAGAGGTTGAGAAGATCAAGAAACGATCAGTGAAGAAATTAAGATGTGTC 422
Db     1817 CTCATCTCTGCAAGAGGTTGAGAAGATCAAGAAACGATCAGTGAAGAAATTAAGATGTGTC 1876
QY     423 AGTAGACATGGAAGAGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGGAGAGGAAT 482
Db     1877 AGTAGACATGGAAGAGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGGAGAGGAAT 1936
QY     483 GATTGCTAGTCTGGGGGGGACTGATTAAGACATATCTGCCAAAGCTGTTCAAGATTAAGCT 542
Db     1937 GATTGCTAGTCTGGGGGGGACTGATTAAGACATATCTGCCAAAGCTGTTCAAGATTAAGCT 1996
QY     543 TGACCAAAATGCTTTTCATTTGGGAGAACATACACACCTGTGGAAAGAGGGAGAGCCAA 602
Db     1997 TGACCAAAATGCTTTTCATTTGGGAGAACATACACACCTGTGGAAAGAGGGAGAGCCAA 2056
QY     603 ACTACTGATGTATGAGAGCTAGCAGAAAGATTCTGTGTGATCATCATGTCATTTGATGAT 662
Db     2057 ACTACTGATGTATGAGAGCTAGCAGAAAGATTCTGTGTGATCATCATGTCATTTGATGAT 2116
QY     663 TACCATTTAAAGATATCTCAAGATTTTCATCCGGGACCTGGAAGATCTCGGAATTCCTTC 722
Db     2117 TACCATTTAAAGATATCTCAAGATTTTCATCCGGGACCTGGAAGATCTCGGAATTCCTTC 2176
QY     723 AGTAGTAAACCAACAGAGAGAGCAGAGACCATTAAGGGAGAAATTAATGATGATCAACA 782
Db     2177 AGTAGTAAACCAACAGAGAGAGCAGAGACCATTAAGGGAGAAATTAATGATGATGATCAACA 2236
QY     783 GAGAGAGCTGATATATGTTATTAACCTAGGTTCTGAACCTCATTTCCGGCATGTGGGAGCC 842
Db     2237 GAGAGAGCTGATATATGTTATTAACCTAGGTTCTGAACCTCATTTCCGGCATGTGGGAGCC 2296
QY     843 TGTATAAACCATTTGTCAAGAGAGTATACATGATTAATTCAGCGATGGGATTTCTTAA 902
Db     2297 TGTATAAACCATTTGTCAAGAGAGTATAGATGATTAATTCAGC--ATGGGATTTCTTAA 2355
QY     903 ATAAAGCTTGGAAAGACCGGATTGACCACTTGAG-----AACATATGACGCTGCCGTCA 956
Db     2356 ATAAAGCTTGGAAAGACCGGATTGACCAACTTGAAGAGGCAATGACGCGCTGCCGTTCAGT 2415
QY     957 GTAACAGATGACTGACAGCGCGTATTTTCTGGTGAATATTTGACAGGTGTAATTCGCTT 1016
Db     2416 ACCAGATGAGACTGACAGCGCGTATTTGACAGGTGTAATTTGACAGGTGTAATTCGCTT 2475
QY    1017 CAATGTCTCCAATTGGAAACGCAATCTCGAAACGTGCAAGCAAGCAATGATGATGTCTAGAG 1076

```

```

Db     2476 CAATGTCTCCAAATTGGAAACAG--ATTCGAAACGTGCAAGCAGAGATTGAAGCTAAAG 2534
QY    1077 CAATTAAGTCTGAGGCTTATCAACAGCAGATAG--AGTGGTACGATGACTCATCAAGCA 1135
Db     2535 CAATTAAGTCTGAGGCTTATCAACAGCAGATAGAAATGAAAGATCTGAATCATCAAGCA 2594
QY    1136 GACGCTTTTGTCAAAGAAAGTAAACAGAGAGAGTGA CAACACACTGTTCTAGACCCAA 1195
Db     2595 GA-GCTTTTGTCAAAGAAAGTAAACAGAGAGAGTGA CAACACACTGTTCAAGACCATT 2653
QY    1196 AATGAGACTGATATTTATAT--GGCTTGCTGCAAGAGAGATATCATCAAGACAGCATTA 1254
Db     2654 AATGAGACTGAAATTTATATATGAGATGCTGAGAGAGAGATCATCAAGACAGCATTA 2713
QY    1255 ACTGAGAGGAGCTCTATTAGCCGAGGAGTCAATCCAACTAGCCCTGATGAGCTCTG-- 1312
Db     2714 ACTGAGAGGAGCTCTATTAGCCCTGAGGAGTCAATCCAACTAGCCCTGATGAGCTCTGAGC 2773
QY    1313 CATGCTGACACACACGAGAGGCTTGTCTAAGTAGAGCAGAAACCTGTTGAGAGAGACCTTA 1372
Db     2774 ATGGCTGACACACACGAGAGGCTTGTCTAAGTAGAGCAGAAACCTGTTGAGAGAGACCTTA 2833
QY    1373 AGCCATTGAATTTGAACCTTGCCAGCATATGCTCTCAAAATGATGATTAAGCCATCA 1432
Db     2834 AGCCATTGAATTTGAACCTTGCCAGCATATGCTCTCAAAATGATGATTAAGCCATCA 2893
QY    1433 GTCCACAGTGAAGCCGTTAATTAAGCAGAAATGATCTAATTGAATCAAGTCAAGTCAAGAG 1492
Db     2894 GTCCACAGTGAAGCCGTTAATTAAGCAGAAATGATCTAATTGAATCAAGTCAAGTCAAGAG 2953
QY    1493 AGAAGCAACCACTTCAAGAACAGCTAGAGGTTTAAATCAACGCTGACAAATGTTT 1552
Db     2954 AGAAGCAACCACTTCAAGAACAGCTAGAGGTTTAAATCAACGCTGACAAATGTTT 3013
QY    1553 GGAAGAAACAGAAACAAAGAGAGCAGACTGATGATGCTGCTGCGCCAGCCAAAGGCTT 1612
Db     3014 GGAAGAAACAGAAACAAAGAGAGCAGACTGATGATGCTGCTGCGCCAGCCAAAGGCTT 3073
QY    1613 CCATGGCGAAATTTGAGAGATTTTGACAGAGTGGCTGATGACAGGAGCGTCATGTTGGC 1672
Db     3074 CCATGGCGAAATTTGAGAGATTTTGACAGAGTGGCTGATGACAGGAGCGTCATGTTGGC 3133
QY    1673 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCAAGAGCAGACTTAATGTCATATGGA 1732
Db     3134 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCAAGAGCAGACTTAATGTCATATGGA 3193
QY    1733 AGTCTGTGCTGCTTTGAAAGCTTAAGAGAAACATTAAGATCTGATGCAAGAAAGCCCA 1792
Db     3194 AGTCTGTGCTGCTTTGAAAGCTTAAGAGAAACATTAAGATCTGATGCAAGAAAGCCCA 3253
QY    1793 GCAGATGCTTGCAGAGATGGCCAAATCTGCAGAGCAAAATATTGACCAAGACAT 1846
Db     3254 GCAGATGCTTGCAGAGATGGCCAAATCTGCAGAGCAAAATATTGACCAAGACAT 3307

RESULT 6
AKI28632      4112 bp      mRNA      linear      PRI 19-FRB-2004
LOCUS        AKI28632
DEFINITION   Homo sapiens cDNA FLJ46791 fis. clone TRACH3029462, moderately
              similar to Actin cross-linking family protein 7.
ACCESSION    AKI28632
VERSION      AKI28632.1 GI:34536103
KEYWORDS     oligo capping; fis (full insert sequence).
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
  Kanohori, K., Ishibaishi, T., Chiba, Y., Fujimoto, K., Hirooka, S.,
  Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotate, T., Watanabe, M.,
  Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Oka, T., Wakamatsu, A.,
  Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hiro, Y., Saito, K.,
  Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M.,

```

Kikuchi, H., Kande, K., Watsushima, M., Takahashi-Fujii, A., Oshima, A.,
Suyiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Negahara, K.,
Masuho, Y., Nagai, K. and Iseogai, T.,
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 4112)
Iseogai, T. and Yamamoto, J.,
Direct Submission
Submitted (15-JUL-2003) Takao Iseogai, FRI Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@kri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction; Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing; HRI and
RAB; annotation: Reverse Proteomics Research Institute, HRI and
RAB.

COMMENT

FEATURES
Source
Location/Qualifiers
1..4112
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TRACH3029462"
/cissue_type="trachea"
/clone_11b="TRACH3"
/note="Cloning vector: PME18SFL3"

ORIGIN

Query Match 80.6%; Score 1637.6; DB 9; Length 4112;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 1778; Conservative 0; Mismatches 58; Indels 18; Gaps 9;

QY 8 GCTGACAGAAACACATCATCTCTCAGC-TCGCCGCCAGCCCTTGAATATGATAC 66
DB 2262 GCTGACAGAAACACATCATCTCTCAGCCTCCGCCGCCCTTGAATATGATAC 2321
QY 67 TCTAAGCAGCAGCAGAGAAACATCGCAACTGCGTATGATAGCTGAACACAGCC 126
DB 2322 TCTAAGCAGCAGCAGAGAAACATCGCAACTGCGTATGATAGCTGAACACAGCC 2381
QY 127 TCATATAGATAGATGAACAAACCTGC---CCACAGTATGATAGACCCCTGGGGAAG 182
DB 2382 TCATATAGATAGATGAACAAACCTGC---CCACAGTATGATAGACCCCTGGGGAAG 2441
QY 183 CTTTCTATCCAGAGAGATATGTGGCAGCCGACACCCCTTACAGTCACTTAAAGAGA 242
DB 2442 CTTTCTATCCAGAGAGATATGTGGCAGCCGACACCCCTTACAGTCACTTAAAGAGA 2501
QY 243 TGTCAAAAAGCGTGTGTGGACCTGGATGAAGCCATTTCTCATCACTCACTTATCA 302
DB 2502 TGTCAAAAAGCGTGTGTGGACCTGGATGAAGCCATTTCTCATCACTCACTTATCA 2561
QY 303 CAAATATAGATAGATCTTGAAGAGCTTGAACGCTGTGGAAGCTCTGAGGACCCACC 362
DB 2562 CAAATATAGATAGATCTTGAAGAGCTTGAACGCTGTGGAAGCTCTGAGGACCCACC 2621
QY 363 CTCTATCTCTGCAAGGTTGAAGAGATCAAGAAACAGATCAAGTAAATTAAGATGTGTC 422
DB 2632 CTCTATCTCTGCAAGGTTGAAGAGATCAAGAAACAGATCAAGTAAATTAAGATGTGTC 2681
QY 423 AGTAGACATGGAAGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGAGAGAAAT 482
DB 2682 AGTAGACATGGAAGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGAGAGAAAT 2741
QY 483 GATTGCTATGATCTGGGGGGAATGATTAAGACATATCTGCCAAGCTGTTCAAGATAGCT 542
DB 2742 GATTGCTATGATCTGGGGGGAATGATTAAGACATATCTGCCAAGCTGTTCAAGATAGCT 2801
QY 543 TGACCAAAATGGTTTTCATTTGGAGAAATATACACACACTGGTGAAGAGAGAGCCAA 602
DB 543 TGACCAAAATGGTTTTCATTTGGAGAAATATACACACACTGGTGAAGAGAGAGCCAA 602

DB 2802 TAACCAAAATGGTTTTCATTTGGAGAAATATACACACACTGGTGAAGAGAGAGCCAA 2861
QY 603 ACTACTGATATGATGAGAGCTAGAGAAAGATTCTGGTGTATCATATGATGATGT 662
DB 2862 ACTACTGATATGATGAGAGCTAGAGAAAGATTCTGGTGTATCATATGATGATGT 2921
QY 663 TACCATTAAGATCTCAAGATTTCAATCCGGGACCTGGAGATCTGGAAATTTGATCTTC 722
DB 2922 TACCATTAAGATCTCAAGATTTCAATCCGGGACCTGGAGATCTGGAAATTTGATCTTC 2981
QY 723 AGTAGTAACCAACAGCAGAGAGAGAGAGACCATTAAGAGAGAAATAGATGACATCA 782
DB 2982 AGTAGTAACCAACAGCAGAGAGAGAGAGACCATTAAGAGAGAAATAGATGACATCA 3041
QY 783 GGAGAGCTGATATAGTTATTAATTAAGTTCTGAACCTCATTTGGCGATGTGGAGACC 842
DB 3042 GGAGAGCTGATATAGTTATTAATTAAGTTCTGAACCTCATTTGGCGATGTGGAGACC 3101
QY 843 TGATTAACCCATTTGCAAGAAAGATATACATGATTAATTAATTCAGCATGGGATCTCTAA 902
DB 3102 TGATTAACCCATTTGCAAGAAAGATATACATGATTAATTAATTCAGCATGGGATCTCTAA 3160
QY 903 ATAAAGCTTGAAGAAACCGAGATTGACCACTTGAG-----AGCAATGCACTGCCGTCA 956
DB 3161 ATAAAGCTTGAAGAAACCGAGATTGACCAACTTGAGAGGCAATGCAAGCTGCCGTCA 3220
QY 957 GTACAGATGACCTGACAGCGGATTTTCTGGGTATATTTGCAAGGTGTATGTTCCGTT 1016
DB 3221 ACCAGAGATGACCTGACAGCGGATTTTCTGGGTATATTTGCAAGGTGTATGTTCCGTT 3280
QY 1017 CAATGCTCCCAATTTGAAACAGCATCTGAAAACCTGCAAGCAGCATTTGATGTCTAGAG 1076
DB 3281 CAATGCTCCCAATTTGAAACAGCATCTGAAAACCTGCAAGCAGCATTTGATGTCTAGAG 3339
QY 1077 CAATTTAAGTCTGAGGCTATCAACAGCAGATAG-AGTGTATGACAGTACTCATCAGCA 1135
DB 3340 CAATTTAAGTCTGAGGCTATCAACAGCAGATAGATTAAGTGAAGATGATATCATCAAGCA 3399
QY 1136 GACGCTTTGCTAAAGAAAGTATACAGNAGAGATGACAAACACCTGTTCTAGACCCAA 1195
DB 3400 GA-CTTTTGTCTAAAGAAAGTATACAGNAGAGATGACAAACACCTGTTCTAGACCCAA 3458
QY 1196 AATGAGCTGATATTTGATAT-AGCTAGCTCGAAGAGATATCATCAACAGCAGATTA 1254
DB 3459 AATGAGCTGATATTTGATATAGGATATCTGGAAGAGAAATCATCAACAGCAGATTA 3518
QY 1255 ACTGAGAGGTCCTCTATTAAGCCTGAGTCAATGCTCCATGATAGTCTCTG- 1312
DB 3519 ACTGAGAGGTCCTCTATTAAGCCTGAGTCAATGCTCCATGATAGTCTCTG- 3578
QY 1313 CATGCTGACACACAGGAGGCTTGAAGTGAAGCAAAACCTGTGGAGAGACCTTAA 1372
DB 3579 ATGCTGACACACAGGAGGCTTGAAGTGAAGCAAAACCTGTGGAGAGACCTTAA 3638
QY 1373 AGCCATTGAATTAAGATCTGCAAGCATCATGCTCTCCAAAATGATATTAAGCCATCA 1432
DB 3639 AGCCATTGAATTAAGATCTGCAAGCATCATGCTCTCCAAAATGATATTAAGCCATCA 3698
QY 1433 GTCCACAGTGAAGCCGTTAAATTAAGCAGAAATGATCTAATTAATCAAGTGAAGAGA 1492
DB 3699 GTCCACAGTGAAGCCGTTAAATTAAGCAGAAATGATCTAATTAATCAAGTGAAGAGA 3758
QY 1493 AGAAGCAAGCAACCTTCAAGCAAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTT 1552
DB 3759 AGAAGCAAGCAACCTTCAAGCAAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTT 3818
QY 1553 GGAAGAAACAGAACAAAGAGAGAGCAGCTGATGATGCTTGGCCCAAGGCAAAAGGTT 1612
DB 3819 GGAAGAAACAGAACAAAGAGAGAGCAGCTGATGATGCTTGGCCCAAGGCAAAAGGTT 3878
QY 1613 CCATGGGGAATTTGAGATTTGACAGAGTGTGATGACAGAGAGGCTCATCTGTTGGC 1672
DB 3879 CCATGGGGAATTTGAGATTTTGGACAGTGTGATGACAGAGAGGCTCATCTGTTGGC 3938

QY	1613	ATCTAAACCCGCTGGAGAGTTTACCGGAAACACGCCAAGACAGACTTAAATGTCATATGGA	1732
Db	3939	ATCTAAACCCGCTGGAGAGTTTACCGGAAACACGCCAAGACAGACTTAAATGTCATATGGA	3998
QY	1733	AGTCGTGCTGCTTTGGAAGCTTAAAGAAAGAAACATATAGATGTCGAGCAAGAAAGCCA	1792
Db	3999	AGTCGTGCTGCTTTGGAAGCTTAAAGAAAGAAACATATAGATGTCGAGCAAGAAAGCCA	4058
QY	1793	GCAGATGCTTGCAAGATGCCCAAAATCTGCAGAGACAAATATATTGACCAAGACAT	1846
Db	4059	GCAGATGCTTGCAAGATGCCCAAAATCTGCAGAGACAAATATATTGACCAAGACAT	4112
RESULT 7			
LOCUS	CQ732923	7194 bp	linear
LOCUS	Sequence	18857 from Patent WO02068579.	PAT 03-FEB-2004
ACCESSION	CQ732923		
VERSION	CQ732923.1	GI:42314338	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Mumukshu; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		
FEATURES	Kits, such as nucleic acid arrays, comprising a majority of		
source	humanexons or transcripts, for detecting expression and other uses		
	thereof		
	Patent: WO 02068579-A 18857 06-SEP-2002;		
	PE Corporation (NY) (US)		
ORIGIN	Location/Qualifiers		
	1..7194		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Query Match	65.2%;	Score 1324.6;	DB 6;
Best Local Similarity	87.0%;	Pred. No. 7.1e-311;	Length 7194;
Matches 1645;	Conservative	0;	Mismatches 54;
		Indels 191;	Gaps 10;
QY	8	GCTGACAGAAACACATCATCATCTTCACAGC-TCGCGCCCGCAGCCCTTGAATATGATAC	66
Db	2580	GCTGACAGAAACACATCATCATCTTCACAGCTTCCGCCCGCAGCCCTTGAATATGATAC	2639
QY	67	TCTAAGCAGACGACGAGAAGAACATCGGCAACTGCGTAGTTGATAGCTGAACACAGCC	126
Db	2640	TCTAAGCAGACGACGAGAAGAACATCGGCAACTGCGTAGTTGATAGCTGAACACAGCC	2699
QY	127	TCATATATATTAAGTAAGTAACAAACTCG-----CCACAGTACTGATAGAGCCCTTGGGGAAG	182
Db	2700	TCATATATATTAAGTAAGTAACAAACTCGGGCCA CAGTTACTGTAATTTGAACCCCTTGGGGAAG	2759
QY	183	CTTTTCTATCCAAGAGATATGTGGCAGCGACACCTTACAGTCACTTAAAGAGA	242
Db	2760	CTTTTCTATCCAAGAGATATGTGGCAGCGACACCTTACAGTCACTTAAAGAGA	2819
QY	243	TGTCAAAAAGCGTGTGTGGCACTGAGTAAGCACTTTCTCAATCACTCAGTTCCATGA	302
Db	2820	TGTCAAAAAGCGTGTGTGGCACTGAGTAAGCACTTTCTCAATCACT-----	2868
QY	303	CAAGTATATCAGATTCCTTGAAGCTTGAAGCACTGTGAAGCTGTGAAGCAGCCACC	362
Db	2869	-----	2868
QY	363	CTCTATCTCTGCAGAGGTTTGAAGATCAA-GAAACAGATCAGTGAATAATAGATGTGT	421
Db	2869	-----CAGAGGTTTGAAGATCAAAGGAACAGATCAGTGAATAATAGATGTGT	2917
QY	422	CAGTAGACATGAAAAAGCTACAGCCGTTGTATGAACTCTTAAACAGAGGGGAGAGAAA	481

D	2918	CAGTAAAGCATGGAAGAAACCTACAGCCCTTGTATGAAACTCTTTAAACAGAGGGAGAGGAA	2977
Q	482	TGATTGCTAGATCTCGGGGGGAGCTGATTAAGACAATATCTGCCAAAGCTGTTCAAGATTAAGC	541
D	2978	TGATTGCTAGATCTGGGGGGGAGCTGATTAAGACAATATCTGCCAAAGCTGTTCAAGATTAAGC	3037
Q	542	TTGACCAAAATGCTTTTCAATTTGGGAGAACATACACACACTGTGTGGAGAGAGAGGAAAGCCA	601
D	3038	TTGACCAAAATGCTTTTCAATTTGGGAGAACATACACACACTGTGTGGAGAGAGAGGAAAGCCA	3097
Q	602	AACACTAGGAGATGTATAGTAGAGCTAGCAGAAAAAGTTCTGTGTGATTCACATCTGCATTTGATG	661
D	3098	AACACTAGGAGATGTATAGTAGAGCTAGCAGAAAAAGTTCTGTGTGATTCACATCTGCATTTGATG	3157
Q	662	TTACCATTTAAAGATACTCAAGATTTCACTCCGGGACCTGGAAGATCTCGAATTAATCTT	721
D	3158	TTACCATTTAAAGATACTCAAGATTTCACTCCGGGACCTGGAAGATCTCGAATTAATCTT	3217
Q	722	CAGTAAATTAACCAACACAGCAAGAACGACAGACCATTAAGGAAAGAAATAGATGACTAC	781
D	3218	CAGTAAATTAACCAACACAGCAAGAACGACAGACCATTAAGGAAAGAAATAGATGACTAC	3277
Q	782	AGGAGGAGCTGGATATATGTTATTAACCTAGGTTCTGAATCTCATTTGGGGGACATGGGGAGC	841
D	3278	AGGAGGAGCTGGATATATGTTATTAACCTAGGTTCTGAATCTCATTTGGGGGACATGGGGAGC	3337
Q	842	CTGATTAACCCATTTGTCAAGAAAGATATATGATGATTAATTAATTCAGCGATGGAGTTCTCTA	901
D	3338	CTGATTAACCCATTTGTCAAGAAAGATATATGATGATTAATTAATTCAGCGATGGAGTTCTCTA	3370
Q	902	AATTAAGCTTGGAAAAGACCGGATTTGACCAACTTGAGAGCAATGCAGCTGCCGTACGTACA	961
D	3371	-----	3370
Q	962	GATGAGACTGCAGCGCGGATTTTCTGGGTAGATATTGACAGTGGTAAGTTCGTTCAATG	1021
D	3371	-----AGCGGATTTTGAAGCTGGTATGATATTTGACAGTGGTAATTAATTAATGCTTCATG	3420
Q	1022	TCTCCAAATTTGGAAACAGCATCTCGAAACTGTCAAGCAGCAGATTTGATGCTTAGAGCAATT	1081
D	3421	TCTCCAAATTTGGAAACAGCATCTCGAAACTGTCAAGCAGCAGATTTGATGAGCTTAAGCAATT	3479
Q	1082	TAACTCTGAGGCTTATCAACAGCAGATAG-AGTGGTACGACTGACTATCACGCAAGCGC	1140
D	3480	TAACTCTGAGGCTTATCAACAGCAGATAGAGAAATGGAAGAACTGAATCATCAACAGCAG-GC	3538
Q	1141	TTTTGCTTAAAGAAAGTAAACAGAGAGGTACAAACAACATGTTCTTAGAACCCAAATAATG	1200
D	3539	TTTTGCTTAAAGAAAGTAAACAGAGAGGTACAAACAACATGTTCTTAGAACCCCAATTAATG	3598
Q	1201	AATCGATATTGATAT-AGCTAGCCTGGAGAGAGAAATCATCAACAGACAGCATTAATCTGG	1259
D	3599	AATCGAATTTGATATGGGATAGCCTGGAGAGAGAAATCATCAACAGACAGCATTAATCTGG	3658
Q	1260	AGGCTGCTTATTAAGCAGTGGGTCAAGTTCCAAACATGCGCTGTATGAGCTCTG--CATGC	1317
D	3659	AGGCTGCTTATTAAGCAGTGGGTCAAGTTCCAAACATGCGCTGTATGAGCTCTGCAATGCG	3718
Q	1318	TGAACACACCCGAGGGCTTGTAAATGAGAGAAACCTGTTGGAGAGAACCCCTTAAGCCA	1377
D	3719	TGAACACACCCGAGGGCTTGTAAATGAGAGAAACCTGTTGGAGAGAACCCCTTAAGCCA	3778
Q	1378	TTGAAATTTGAACCTTGCACAGCATCATAGTGTCCAAATATGATTAATAGCCCATCAGTCCA	1437
D	3779	TTGAAATTTGAACCTTGCACAGCATCATAGTGTCCAAATATGATTAATAGCCCATCAGTCCA	3838
Q	1438	CAGTGAAGCCGTTTAATTAAGCAGAGAAATGATCTAAATTTGATCAAGTGCAGAGAGAGAG	1497
D	3839	CAGTGAAGCCGTTTAATTAAGCAGAGAAATGATCTAAATTTGATCAAGTGCAGAGAGAGAG	3898
Q	1498	CAAGCAACTTTCAGAACAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAA	1557
D	3899	CAAGCAACTTTCAGAACAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAA	3958

ARKRAQFHEANSKIMELFESEKSLDSELEIANDPDKIKAOLVOKHEFOKSLGKHS
VDTNRGSLKETS LADNLIKLDNMLSELKWDITCGKVEONKLEBELTSG
OFDIALQALIDMLRYVEPOLAEDQPVHGIDIDVMSLIDNKHQVEKSLTSVQK
RSARELEIGSRDSDSWRVOMELSTRMEVCAISJSGKORLEKQAEBSVYHT
LMELEABOTLBFHGLDDEDELRTLIEHKEPKRLEEKRAELSKATQEDALLA
VCHPDSITTTIKMITTIQARFEVILAMAKHOORLAGLAGLIKABELLETLAMICW
AETTLKEDKEVIPOSEBEVKULIAHQFMEBMTKOPDVDTKYRRAVDPSL
OSHI PVLIDGKRAKRF PASGFYPSQSOTIETKNRVMVLVSKWOQVMLALERK
LNDALRLBELREFANFDPDIWRKKTMRWNHKS RVMPFRF1DIDOGKITROEPI
DGLISKEFPLSRLEMSAVADI FDRDGDYDIYEFVALAHPENDAYKPLTDADKIGDE
VTRQVAKCKAKRFOYBOJGDNKRFELNGPDGSOOLRVITLSESTWVRVGGGMA
LBERLYKNDPCRVHHGSKMLBESNSSTIATOPTAKRTIMELBKPLIADGACOG
MAAFPRGRSRPSRSGASPNKSTSSSHACQNASPVPAASVTYKGTPIQSKLPL
GYSJGGFHSGEDSALITTAARVROFESRTPSRPGSRASKASRSRSGDA
SPDISEIOSVCSDETVPTQHRPVRPAGSRPSTAKPSKIPTRQRRSPASKLKSSKR
"

ORIGIN

Query Match 58.6%; Score 1190.8; DB 10; Length 17206;
Best Local Similarity 80.3%; Pred. No. 2.6e-278;
Matches 1507; Conservative 0; Mismatches 351; Indels 18; Gaps 9;

QY 4 CCATGCTGACAGAAACCAATCATCTCTGAGCT-CCCGGCCGACCCCTTGAAATANG 62
DB 11291 CTTGGCTTAACGGAAACCAACGATCATCTCTGAGCTGCTGCGCCGACCTTGAGATAG 11350
QY 63 ATACTTAAGGCGACGACGAGAGAAACATCGGCACTGCGTAGATTGATAGTAAACACA 122
DB 11351 AGACTTTTGAAGACGACGAGAGAGACCGGCACTTGAGAACTGATAGCCGACACA 11410
QY 123 AGCTCATATAGTAAAGTGAACAAACTGG---CCACGTACTGATAGCCCTGGGG 178
DB 11411 AGCCACACATGATGATGATGACAAAGCTGGGCGACAGCTATTGGAATGAGCCCAAG 11470
QY 179 AAGCTTTTCTATCCAAAGAGATATGTGGCAGCCGACACCCCTTACAGTCAATTAAG 238
DB 11471 AAGGCTTTTATTCAGAAAGATATGTGGCAGCCGACACCCCTTACAGTCAATTAAG 11530
QY 239 AAGATGTCAAAAACGCTGTGTGGCACTGGATGAGCCATTTCTCAATCAACTCAGTTC 298
DB 11531 AAGATGTCAAAAACGAGCTGTGTCTTGA TGAAGGCATTTCTCAGTCAACCAAGTTTC 11590
QY 299 ATGACAGATAGTCAAGTCTTGAAGCCTGGACGCGATCGTGAACGCTTGAAGCGAC 358
DB 11591 ATGACAGATGCGACGAGATCTTGAAGCCTGGACGCGATCGTGAAGCCTTGAAGCGAC 11650
QY 359 CACCTCTATCTGTGACAGGTTGAGAAATCAAGAAACGATCAGTGAATAATTAAGT 418
DB 11651 CGCCATCTATCTCGCGAGGTGAGAAATCAAGAAACGATTTGTAATTAAGGCG 11710
QY 419 TGTCACTAGACATGAGAAAGCTACAGCCGTTGTATGAACTCTTAAACAGAGGGAGAG 478
DB 11711 TGTCCGTGACATGAGAAAGCTCAGCCGCTATACAGACTCTGAGGACAGAGGGAGAG 11770
QY 479 AATGATTTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAAGCTGTTCAAGAT 538
DB 11771 AGATGATCGCAGATCTGAGGGGACCTGAAACGATATCTGCAGAGCTGTACAGAT 11830
QY 539 AGCTTACCAAAATGTTTTCAATTTGGGAGAACATACACACATGCTGAGGAGAGAGAG 598
DB 11831 AGTTGACCAAAATGTTTTCAATCTGGGAGAACATCCACACGCTGTGAGAAAGCGGAG 11890
QY 599 CCAAACTAGTGTATGAGTGTAGCTAGCAGAAAAAGTTCTGGTGTGATCAATGCTCATTTGA 658
DB 11891 CCAAACTAGTGTATGAGTGTAGCTAGCAGAAAAATTTCTGGTGTGATCAATGCTCATTTGA 11950
QY 659 TAGTTACCATTTAAAGTACTCAAGATTTCA TCCGGGACCTGGAAGATCTGGAATTGATC 718
DB 11951 TCTCTCAATCAATCAAGACACTCAAGATTTCA TCCGGGACCTGGAAGATCTGGAATTGATC 12010
QY 719 CTTCACTAGTAAACAAACAGCAGAGACGACAGACCAATTAAGAGAAATGATGAGAC 778
DB 12011 CTTGGGTGTAAGACAGCAGAGAAAGCTGCAAGGCAATCAAGAGAGAAATTGACGGCC 12070

QY 779 TACAGAGAGAGCTGATATATATATTAACCTAGTTCGAACTCATTTGGCGCATGTGGG 838
DB 12071 TGAAGAGAGAGCTGACATGTGATCATACGCTAGGCTCAGAGCTCATCGGTGATGTGGG 12130
QY 839 AGCTGATTAACCATTTGTCAGAGAGATATACATGATTAATTAGCGATGGGATTTCT 898
DB 12131 AGCTTACAAACCATTTGTCAGAGAGATATGACATGACATTAATTGACG-ATGGGATTTCT 12189
QY 899 CTAATAAAGCTTGGAAAGCCGGATGTAACAATTGA-----GAGCAATGAGCTGCC 952
DB 12190 CTAATAAAGCTTGGAAAGCCGGATGTAACAAGCTGTAAGAGGCCATGCAAGCTGCCGT 12249
QY 953 GTCACTAGACATGAGCTGACAGCGGATTTTCTGGGTAGATATTTGACAGGTGTAAGTTC 1012
DB 12250 CAGTACAGAGAGCGAATTGACAGGGAATATTGACTGGGTGATGATTCGACAGCGCAAGTTA 12309
QY 1013 GCTTCAATGTCTTCATTTGGAACAGCATCTCGAAACTGTTCACAGACAGATTAATGTCT 1072
DB 12310 GCCACATGTCTCCCATTTGGAACAG-ACCTGAGAGACCTTCAGACAGATTTGMAAGCT 12368
QY 1073 AGACCAATTTAAGTCTGAGCGCTATCAACAGCAGATAG-AGTGTGACACTGACTATCA 1131
DB 12369 GAAACATTTTAAGTCTGAAAGCTTATCAACACAGATGAAATGAAAGGCTGAACCATCA 12428
QY 1132 CGCAGACGCTTTTCTTAAAGAAAGTAAACAGNAGAGATGACAAACACTGTTCTAGACC 1191
DB 12429 AGCAGA-ACGTGTTGTAAGAAAGTAAACAGAGAGCAGACACACAGTCCAGAGACC 12487
QY 1192 CAAAAATGAACTGATATTGATNT-CCGTAAGCTGGAAGAGAAATCATACACACAGC 1250
DB 12488 CGCTGATGAGCTGAAGTGAATCTGGAACAGCTTGACAGGCAATTTGACAGACAGC 12547
QY 1251 ATAAATGAGAGGTGCTTATTAAGCCNRTGGTCAAGTTCACAACTGCTGATGAGCTCC 1310
DB 12548 ACAAGCTCAAGGTGCTCTTTAGCACTGGGTCAAGTTCACAGACGCTTGAGAGACTCC 12607
QY 1311 T-GCATGCTGACACACACCGAGGGGCTTCTAAGTGAGAGAAACCTGTGAGAGAGACC 1368
DB 12608 TGGCATGTGTCACACACCAAAAGCTTGCTAAGTAAACAGAAACCGTTGGAGAGAGACC 12667
QY 1369 CTAAAGCCATTGAATTTGAATTTGCCAGACATCATGTCTCCAAAATGATATTAGCCC 1428
DB 12668 CGAAAGCCATTGAATTTGAATTTGGCTAAGCATCATGTCTCCAAAACGAGTCTTAGCCC 12727
QY 1429 ATCAGTCCACAGTGAAGCGCTTAATTAAGCAGAAATGATCTAATTTGAATCAAGTGCAG 1488
DB 12728 ACCAGTCCACGTTAGAGCCGTTACAAAGCAGAAATGATCTCATGAGTGCAGATGAG 12787
QY 1489 GAGAAAGCAAGCAACCTTCAAGAACAGCTAGAGGTTTTTAATCAACGCTGCAAAATG 1548
DB 12788 GGAAGAAACCAAGCAACCTTCAAGAACAGCTAGAGGTTTTTAATCAACGCTGCAAGAT 12847
QY 1549 TTTTGAAGAAACAGAAACAAAGAGCAGCAGCTGATGTGCTTCCAGGCGCAAG 1608
DB 12848 TTTTGAAGAAACAGATCAAAAGAAAGCAGCTGACAGCCCTTCCAGGCGCAAG 12907
QY 1609 GGTTCATGCGGAAATTTGAGATTTGACAGAGTGTGCTGACACAGAGCGTCACTGT 1668
DB 12908 GGTTCACGCTGAATTCGAAGATTTGACAGAGTGTGACAGACAGAGCGTCACTGT 12967
QY 1669 TGGCATCTAAACGCTGGAGGTTTTACCGGAAACAGCCAGAGAGCAGTTAAATGTCATAT 1728
DB 12968 TGGCATCTAAACCTCTGGAGGTTCTCCAGAGACAGCCAGAGAGAGCTTAAACGCAATAT 13027
QY 1729 TGGAGTCTGTGCTGCTTTGAAGCTTAAGAAAGAAACATTAAGATCTGATGACAGAA 1788
DB 13028 TGGAGTCTGACCGGCTTTGCAATCAAGAAAGAGAGCTGAATGCTGATGAGG 13087
QY 1789 GCACAGATGCTTGAAGATCCCAAAATCTGACAGAGCAATATTTGACCAAGATATTA 1848
DB 13088 GGCAGAGATGCTTGAAGATGTTCCCGATCTGCGAGAGCAACATTTGACCAAGATATTA 13147

LDKLVKDYDAVQVINSERHLQLERAQSLVSQFWETYEELIEMWLTETORIIISQUPAPAL
EYETFERQOEERHQLRELJAEHPHIDIKMKTGPOLLESPKGIYIOEKYVADTLY
SOITEDVSKRAVVLDEAISOSTOFHDKIDITLESRLIARQPSISAEVYKIEO
IGENKSVYDMEKLOPYLETILROGEBEMILARSGETKDEARAVODKIDMSWIMESI
HILYERERAKLDIMELAEKFWCDHMSLVYITKDQDFIRIDLEDPSGIDBSVAKQOEA
AEAIREEIDGLOEBELHNVITLISELTAACEEPDKPIVKSIDELNSAMBSLNAKMDR
VDLEBAMQAAVQYODGLOGIFDWMVJAGKLTAMSPITGDTLETVQOILQIELQFKE
AYOQOMERMLNHOAEILKLKYTEADKIVODPLMELKJIMSLIBERIVSRHOKLEG
TLTALQFOHALDELLAMLTHTKGLSECPGDPKALIEIARKHVIONDPLAOKS
TYEAVNKAQNDLIESSEGEASNLQYKRLINORWODIJEKTORCOQDLSAROKS
FHEIEEDLOOMLDTBRHLIASKPLGSLPETAKEQJAMHEVTTARILKEETKSLML
RQOQMLARCPRSASATNI DDITNLKEKESVSKLNEKTKLEALHLAMFNINSIQD
FINMLTQAGTCLNVAARPSLIDITLIFQIDENKRVFNENVSHQOIIEDKCTHLKY
FSQKQDVVILIKMLISVOSKEWKEVORLEBERAALQKCFEASKMEIEMER
SEKSLDESELEIANDPKIKAOLOVHKEFOFSGKSVYDTTRTSRLEKSLSLD
NKLDMMLSELBDMPTIGCSYVERONKLEALLFSGOFTDALQALIDMYREROLA
BOQPVHGDIDIVYSLIDNHKVPKEKEGRTSSVQALKRARELIESRDSBVSATVQM
QELSTRVETVCALISIKQTRLESALQOAEFHSVHTLLEWLEAEQTLRPHGALPDD
EDALRTLBOHKEFMRLEBEKRAELSKAETGMGDALLAVCHPDSITTKIMITIIQAR
EVLAMAKHOQOQLAGLAGLAKOELLETLAMLOWAEETLEKDEKREVIPOIIEBK
TLIAEHOTFMEEMTRKOPDVDTKYTKRRATDPGLQSHIPLYDKGRAGKRPASG
FYPSGOTOIETKNPRVNLVSKMOQWMLAERRKNDLALRELEIPEFPAEPDI
WRKRYVRMMNHKRSVMDPFRRIDKQDQKITRQERTIDGLSKPFTSRLEMSAVADI
FDRDGYIDYIEFVAALHPNKDAYRPIVDADKIGDENVKQVAKCFACARFOYEQIGD
NKRFPLGNQFQDSQOQLRVLRIKSTVMVRVGGMMALDEFLVKNDCPRCVHHSKML
RSEKNSITATQPTLAKRTMELREKFIADAGSQGMAAFRGRGRSPRSKASPN
RSTGSSHACQASPVPVPAASPTPKGPIQSGSKRLPGVYSGSGESGSESLITPA
ARYRQFASBRKTPRSGSRGSKASGRASRSRGSDASDPDISIUSVCSDEVTVEQT
HREVPAGSRPSTAKPSKITTPQRRSPASKLDSKRG"

ORIGIN

Query Match 58.6%; Score 1189.2; DB 10; Length 22218;
Best Local Similarity 80.3%; Pred. No. 6.6e-278;
Matches 1506; Conservative 0; Mismatches 352; Indels 18; Gaps 9;

QY 4 CCATGCTGACAAACCAATCAATCACTCTCAAGT-CCGCGCCCAAGCCCTTGAATANG 62
DB 17318 CCGGCTAAAGGAAACCAACGAATCATCTCTAGCTGCGTGCAGCCCTTGAGATAGC 17377
QY 63 ATACTCTAAGGCGACGAGAGAAACATCGGCAACTGCGTGTGATGTCGTAACACAC 122
DB 17378 AGACTTTTGAGACAGCAAGAGAGACCGGCAACTTTCGAGAACTGATAGCGGAGACAC 17437
QY 123 AGCCTATATAGATAAGATGAAACAAAACCTGG---CCACAGTACTGAATAGCCCTTGAGG 178
DB 17438 AGCACAACATAGATTAGATGACAAAGCTGGCCACAGCTATTGGAATGAGCCCAAAG 17497
QY 179 AAGGCTTTTCTATCCAAAGAGATGTGGCAAGCCGACACCCCTTTTACAGTCACATTAAAG 228
DB 17498 AAGGCAATTTATATCAAGAAAGATATGTGGCAGCCGACACCTCTACAGTCAAAATTAAAG 17557
QY 239 AAGATGCAAAAAGCGGTGCTGTGGCACTGATGTAAGCCATTTTCTCAATCAACTCACTTC 298
DB 17558 AAGATGCAAAAAGCGAGCTGTGGTCTTGTATGAGGCAATTTTCTCACTCAACCCAGTTTC 17617
QY 299 ATGACAAGATAGATCGATCCTTGAAGAGCTTGAAGCGCATCGTGAAGCGTCTGAGGAGC 358
DB 17618 ATGACAAGATCGACAGATCTTGAAGAGCTTGAAGCGCATTTGGGAGCGTCTCGGAGC 17677
QY 359 CACCTCTATCTGTGAGAGGTTGAGAAAGATCAAGAAACAGATCACTGAAATTAAGATG 418
DB 17678 CGGCATTCATCTCCGCGAGAGGTGAGAAAGATCAAGAGGAGATTTGGTGAATAATAGAGG 17737
QY 419 TGTCAATAGACAGAGAAAGCTACAGCCGTTGTATGATAACTCTTAAACAGAGGGAGAGG 478
DB 17738 TGTCCGTGACATGAGAAAGCTCAGCCGCTATATACAGACTCTGAGGCAAGAGGAGAGG 17797
QY 479 AATGATTTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCAAAAGCTTTCAGAGAT 538
DB 17798 AGATGATCCGCAATCCGAGGGACATGAAAAAGCTATCTGCAAGACTGTACAGAGATA 17857
QY 539 AGCTTGAACAAATGGTTTTCATTGGAGAAATACACACACTGTGTAAGAGAGGAGAG 598
DB 17857 AGCTTGAACAAATGGTTTTCATTGGAGAAATACACACACTGTGTAAGAGAGGAGAG 598

DB 17858 AGTTGACCAAAATGTTTTCATCTGGGAGAGCATCCACGCTGTCGAGAAACGGGAG 17917
QY 599 CCAAACTCACTGATGTGATGAGAGCTAGCAGAAAAGTTGTGTGTGATCACTATCATTTGA 658
DB 17918 CCAAACTGCTGATGTGATGAGAGCTAGCAGAAAATTTCTGTGTGATCACTATCATTTG 17977
QY 659 TAGTTACATTAAGATTAAGATTAAGATTTTATCGGAGACCTGGAAGATCTGGAATTAATC 718
DB 17978 TGTCACTCACTCAAAAGACATCAAGATTTTATCTCGGAGACCTGGAAGATCTCGGAGAT 18037
QY 719 CTTCAGTAGTAAACACACGCAAGAAAGCAGCAGACCTTAAGGAGAAATATGATGAC 778
DB 18038 CTTGCTGCTTAAAGCGCAGCAAGAAAGCTGACAGGCAATCAAGAGGAGATTAAGCGGCC 18097
QY 779 TACAGAGAGAGCTGATATAGTTATTAACCTTGATTTCTGAACCTATGGGACATGNGGG 838
DB 18098 TGCAGAGAGAGCTGACATGATCATACCTGAGCTCAAGCTCATCGCTGCAATGNGGG 18157
QY 839 AGCTGATTAACCCATTTGTCAAGAAAGTATACATGATTTAAATTCAGGATGGAATTC 898
DB 18158 AGCTTGACAAACCATTTGTCAAGAAAGCATTTGACGAGCTTAATTCAGC-ATGGATTTCT 18216
QY 899 CTAAATTAAGCTTGAAGAACCGGATTCACAACTTGA-----GAGCAATGCACTGCC 952
DB 18217 CTAAATTAAGCTTGAAGAACCGGATTCACAACTTGA-----GAGCAATGCACTGCC 18276
QY 953 GTCAGTACAGATGACATGACGCGGATTTTCTGGGTGATATTTGACAGTGGTATGTTCC 1012
DB 18277 CAGTACAGAGAGATTTGACGAGAAATATTCAGCTGGGTGATCATCGAGCGCAAGTTA 18336
QY 1013 GCTTCATGTCCTCAATTTGGAACAGCATCTGAAACTGTCAAGCAGCAGATTTGATGCT 1072
DB 18337 GCACAGATGTCCTCCATTGGAACAG-ACCTGAGACCGTCAAGCAGCAGATTTGAAGAGCT 18395
QY 1073 AGAGCAATTTAAGTCTGAGGCTATTAACAGCAGATAG-AGTGTATGCACTGACTATCA 1131
DB 18396 GAAAGCAATTTAAGTCAAGAGCTTATTCMAACAGATGAAATGGAAGGCTTGAACCATCA 18455
QY 1132 CGCAGAGCTTTTGTCTAAAGAAAGTAAAGNAGAGTGAACAACACATGTTCTAGACC 1191
DB 18456 AGCAG-ACCTGTGCTGAAGAAAGTAAAGNAGAGTGAACAACACATGTTCTAGACC 18514
QY 1192 CAAAAATGANTGATATTGATAT--GGCTAGCTGAGNAGAGATCATCAACAGACAG 1250
DB 18515 CGCTAGTGAAGCTGAAGTGTGATCTGGGACAGCTGGAACAGGAATTTGACAGACAG 18574
QY 1251 ATTAACCTGAGAGGTGCTTATTAGCCNTGGGTCAAGTTCCAAACATGCTCTGTATGAGCTCC 1310
DB 18575 ACAAGCTCAGAGGTACTCTTATAGCACTGGGTCAAGTTCCAGACAGCCCTGAGCAGAGCTCC 18634
QY 1311 T--GCATGTGACACACACCGAGGCGTGTCTAAGAGAGAAACCTGTGTGAGAGAGACC 1368
DB 18635 TCGCATGTGTGACACACACCAAGGCTGTCTAAGTGAAGAAACCGTTGAGAGAGACC 18694
QY 1369 CTAAAGCCATTAAGATTTGAATTTGCAAGCATCATATGCTCCAAAATGATGATTAGGCC 1428
DB 18695 CCAAGCCATTAAGATTTGAATTTGCTGTAAGCATCACTGCTCCAAAACGACGCTTAGGCC 18754
QY 1429 ATCAGTCAACAGTGAAGGCGTTAATTAAGCAGAAATAGATCTTAATTAATCAAGTGACG 1488
DB 18755 ACCAGTCAACAGTGAAGGCGTTAATTAAGCAGAAATAGATCTTAATTAATCAAGTGACG 18814
QY 1489 GAGAAAGACCAAGCACTTTCGAAACACAGTGAAGTTTAAATCAACGCTGCAAAAG 1548
DB 18815 GCAAGAAACCAAGCACTTTCGAAACACAGTGAAGTTTAAATCAACGCTGCAAAAG 18874
QY 1549 TTTTGAAGAAACAGAACAAAGAAAGCAGACCTGAGTGTGCTTCCAGGCGCAAG 1608
DB 18875 TTTTGAAGAAACAGAACAAAGAAAGCAGACCTGAGTGTGCTTCCAGGCGCAAG 18934
QY 1609 GGTTCATGAGCAATTTGAAGATTTGACAGATGCTGACATGACAGAGCGTCACTGT 1668
DB 18935 GGTTCATGAGCAATTTGAAGATTTGACAGATGCTGACAGAGCGTCACTGT 18994

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
AUTHORS	1 Venter,C.J.,Adams,M.C., Li,P.W. and Myers,E.W.									
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcript, for detecting expression and other uses thereof									
JOURNAL	Patent: WO 02068579-A 15796 06-SEP-2002;									
FEATURES	PE Corporation (NY) (US)									
Source	Location/Qualifiers									
	1..15293									
	/organism="Homo sapiens"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:9606"									
ORIGIN										
Query Match	33.2%;	Score 674.4;	DB 6;	Length 15293;						
Best Local Similarity	63.2%;	Pred. No. 9.3e-153;								
Matches 1134;	Conservative 0;	Mismatches 645;	Indels 15;	Gaps 6;						
QY	41	CCCCCCCCCCTTGATATGATTACTTAAGCAGACGAGAGAGACATCGGCAACTG	100							
DB	10554	CCCTTCACGCCATTGATGATGAGAGCTCAGCAGACAGAGAGAGAGAGATTA	10613							
QY	101	CGTAGTTGATGAGCTGAACCAAGCTCATATGATGATGAACTGGCC---A	156							
DB	10614	AGGGAATCTATTGCTGTAACCAACCTCATATGACAACTACTAAAGATAGGCCACA	10673							
QY	157	CAGTACTGAATGAGCCCTGGGGAGGCTTTCTATCCAGAGAGATGAGAGCCGAC	216							
DB	10674	CTAAAGGATTAACCCCTGAGAGAGGGGAAATGGTGAAGAAATTCAGAAAGCGAA	10733							
QY	217	ACCCTTACAGTCACATTAAGAAGATGTCAAAAAGCGTCTGTGCACTGATGAAGCC	276							
DB	10734	AACATGTATGCCCAATTAAGAGAGAGTCCGCCAGAGAGCCCTGGCTGTGATGAAGCC	10793							
QY	277	ATTTCATCACTCACTCAGTTCCATGACCAAGATGATCAGTCCCTGAGAGCTTGAAGCGC	336							
DB	10794	GTCGCCAGTCCACACAGTTTCATGATTAATGAGCTATGTTGAGACACTGAGAAAT	10853							
QY	337	ATCGTGAAGCGTGTGAGGAGCAGCACCCTCATCTCTGAGAGAGTGAAGATCAAGAA	396							
DB	10854	CTTTCTCTGCGCTGCGGTATGCCACCATGATCCCTGTGAATTAACAAGATCAAGAG	10913							
QY	397	CAGATCAGTGAAGATTAAGATGTGTCAATGACATGAGAGAGAGAGAGAGTGA	456							
DB	10914	TGATCAGTGAACATTAAGATGTGTCAACCGTGAGACTGAAGAAATGTGACGCATCTTTGAG	10973							
QY	457	ACTCTTAACAGAGGGGAGAGAAATGATTGTAAGATCTGGGGGAGCTGATTAAGCATA	516							
DB	10974	GCTTGAAGCGCGTGAGAGAGACTTATTTGACGATCTCAGGAGACAGCAAGATCTTG	11033							
QY	517	TTCGCCAAGCTGTTCCAGATAAGCTTGACCAATGATGTTTCATTTGGGAGAACTACAC	576							
DB	11034	GCTGCAAAAGAAATCCAGATTAATTTGATCAATATGATTTCTTCTGGAGAGATCAAA	11093							
QY	577	ACAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	636							
DB	11094	GCTCGGGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	11153							
QY	637	TGCTGATGATCACTGTTCATGATTAAGTTACATTAAGATTAAGTATTAAGTTTCAATCGGGAG	696							
DB	11154	TGCTGATGATCACTGAGCTCTCCAGACCATCAATAAAGACCCAGAGATATTGTCATGAC	11213							
QY	697	CTGGAAGATCTGGAATTTGATCTCTGATGATTAACCAAGAGAGAGAGAGAGAGAG	756							
DB	11214	TTGGAAGAGCCAGGATTTGATCTCTTCATATATCAACACAGGTTGAGCTGTGAGACT	11273							
QY	757	ATAAGGAGAGAAATGATGAGACTTACAGAGAGAGCTGATATATGTTATTAACCTGAGTTCT	816							
DB	11274	ATTAAAGGAGAGACAGATGTGCTGCATGAAGAGACTGGAAGTTATTTGGATCTTGGAGGA	11333							
QY	817	GAACTCAATTCGGGAGTGGGGAGCTGATTAACCATTTGTCAAGAGAGATACATGAG	876							

DB	11334	GATTTGATTTTTGCTGTGAGAGAACTGAGAACTGAAAGTGAAGAGATGATGATGAG	11393			
QY	877	TTAAATTCAGCGATGAGATTTCTTAATTAAGCTTGAAGAGAGAGATGAGAGATGAG	936			
DB	11394	ATCAATTAATGC-TTGGAGAGACTTAACAAACATGAGAGAGAGAGAGAGAGAGAG	11452			
QY	937	G-----AGCAATGACGCTCCCTCAGTACAGATGAGCTGAGCGGATTTTTCTGGGT	990			
DB	11453	GATGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	11512			
QY	991	AGATATTCAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1050			
DB	11513	AGATTAACATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	11571			
QY	1051	TCAAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1110			
DB	11572	TTAAAGTACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	11631			
QY	1111	AGTGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1170			
DB	11632	AGATGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	11691			
QY	1171	ACAAACACCTGTTTGAAGACCCAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1229			
DB	11692	ACAGAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	11751			
QY	1230	AGAGATCAATCAACAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1289			
DB	11752	AGAAATGCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	11811			
QY	1290	AACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1347			
DB	11812	AGATGCTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	11871			
QY	1348	AGAACTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1407			
DB	11872	AGAGACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	11931			
QY	1408	TCCAAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1467			
DB	11932	TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	11991			
QY	1468	ATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1527			
DB	11992	AGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	12051			
QY	1528	TAAATCAAGCTGAGCAAAATGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1587			
DB	12052	TGAACCAATGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	12111			
QY	1588	GTCCTTGGGCGAGGCAAGAGGCTCATGGGAAATTAAGATGATGATGATGATGATGATGAT	1647			
DB	12112	CAACTCTGAGAGAGGCGCCAGGCTTCAAGTGAATTAATTAATTAATTAATTAATTAATTA	12171			
QY	1648	CTGACAGAGAGGCTCATGTTGTCATCTTAACCGCTGGAGGTTTACCGGAAACAGCA	1707			
DB	12172	CTAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	12231			
QY	1708	AGAGAGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1767			
DB	12232	GAGAGAGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	12291			
QY	1768	ATAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1821			
DB	12292	ATAATCACTACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	12345			

RESULT 12
AX677291 6007 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 69 from Patent WO02086122.
ACCESSION AX677291
VERSION AX677291.1 GI:29334696

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1

AUTHORS Legrain, P. and Daviet, L.

TITLE Protein-protein interactions in adipocytes

JOURNAL Patent: WO 02086122-A 69 31-OCT-2002;

Hydrigenics (FR)

FEATURES

Location/Qualifiers

Source 1..6007

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 32.3%; Score 655.4; DB 6; Length 6007;

Best Local Similarity 62.9%; Pred. No. 3.5e-148;

Matches 1134; Conservative 0; Mismatches 645; Indels 24; Gaps 7;

QY 41 CCCCCCCCCCTTGAATGATCTTAAGGCGACGAGAAACATCGCACTG 100

DB 1259 CCCCTCCAGCATTTGATGATGAGAGCTCAGGCGCAACAGAGAAATGAGGCAATTA 1318

QY 101 CGTAGTTGATGATGAGACACAGGCTCATATAGATTAAGTAAACAAACCTGGCC---A 156

DB 1319 AGGAAATCTATTGCTGAGACAAACCTCATATGCAAACTAATTAAGATAGGCCACAA 1378

QY 157 CAGTACTGATGAGAGCCCTGGGAAAGCTTTCTATCCAAAGAAAGTATGTGGACCCAG 216

DB 1379 CTAAGAGATTAACCCCTGAGGAAAGGGAATGTGAGAAATAACAGAAAGAGAA 1438

QY 217 ACCCTTTCAGTCACTTAATAAGAAATGTCAAAAAGCTGTGGCACTGGATGAAGCC 276

DB 1439 AACATGATGCCCCAAATTAAGAGAGGTGGCCGACGAGCCCTGGCTCTGATGAAGCC 1498

QY 277 ATTCTCAATCAATC-----AGTTTCATGACATAGATGATGATGCTGGAGAGC 327

DB 1499 GTGTCCAGTCCACACAGATTAACAGAGTTTCAATGATTAATTAAGATGATGAGCA 1558

QY 328 CTGAGAGCATCTGAGAACTCTGAGGAGCCACCTCTATCTGACAGAGTTGAGAG 387

DB 1559 CTGAGAGATCTTCTCTGCTGCTGCTGATGACCACTGATCCCTGCTGAGATGAGCAAG 1618

QY 388 ATCAAGAAACAGATCAGTGAATAATGAGATGTCTCAGTGAACATGAGAAAGCTACAGCG 447

DB 1619 ATCAGAGGTGATCAGTGAACATTAAGAGTCCACCGTGGAGCTAAGAAACTGCAAGCA 1678

QY 448 TTGTATGAACTTTAAACAGAGGAGAGAAATGATGCTAGATCTGGGGAGATGAT 507

DB 1679 TCTTTGAGGCTTTGAGAGCCCGTGGAGAGAGCTTATGAGACATCTCAGAGAGCAAC 1728

QY 508 AAAGACATATCTGCCAAAGCTGTTACAGGATAGCTTGACCAATGTTTTCATTGAGAG 567

DB 1739 AAGATCTGCTGCAAAAGAAATCCAGAGTAAATTTGATCAATGTGATTTCTTCTGGAG 1798

QY 568 AACATACACACTGTGTGAGAGAGAGGAGCCAAACTACTGATGTGATGAGAGCTAGCA 627

DB 1799 GACATCAAGCTCGGGCTGAGAGAAAGAAATCAAAATTTCTTGATGTCTTGAATTAGCA 1858

QY 628 GAAAGTTCTGAGTGTGATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 687

DB 1859 GAGAGTTCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1918

QY 688 ATCCGGGACCTGAGAGATCTGGAATGATCTTTCAGTGTGATGATGATGATGATGATGAT 747

DB 1919 GTCTATGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1978

QY 748 GCAGAGACATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 807

DB 1979 GCTGAGACTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2038

QY 808 CTAGGTTCTGAATCAATTCGCGCATGTGGGAGCCGTGATTAACCATTTGTCAGAAAGAT 867

DB 2039 CTTGAGAGAGATTTGATTTTCTGCTGTGAGAGAACTGAGAGAGCTGAGAGAGAGAG 2098

QY 868 ATACATGATTTAAATTCAGCGATGGATTTCTTAATTAAGCTTGGAAAGACCGGATGGA 927

DB 2099 ATTGATGAGATGAATTAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2157

QY 928 CCAACTTGAAG-----AGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 981

DB 2158 AAAACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2217

QY 982 TTTCTGGAGATATTCGAGAGGTGTTAAGTTGCTTCAATGCTTCAATTTGAGAGAGATC 1041

DB 2218 TGACTGCTGATGATTAACACTGTGATTAATCTGACCATGCTGCTGCTGCTGCTGCTG 2276

QY 1042 TCGAACTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1101

DB 2277 TCAATATCTGTTAAGATCAAGTTAATGAATGAAGAGAGAGAGAGAGAGAGAGAGAG 2336

QY 1102 AGCAGATGAGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1161

DB 2337 AGCAAAATTTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2396

QY 1162 NAGAGAGTGAACAAACACACTGTTCTAGACCCAAATAGAGAGAGAGAGAGAGAGAGAG 1220

DB 2397 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2456

QY 1221 GCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1280

DB 2457 ACCTGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2516

QY 1281 GTCACTTCCAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1338

DB 2517 GTCACTTCCAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2576

QY 1339 TAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1398

DB 2577 TGAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2636

QY 1399 ATCATGTGCTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1458

DB 2637 ACATGTGCTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2696

QY 1459 CAGAAATGATCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1518

DB 2697 CTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2756

QY 1519 TAGAGTTTAAATCAACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1578

DB 2757 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2816

QY 1579 AGCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1638

DB 2817 AGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2876

QY 1639 AGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1698

DB 2877 TGAAGCTTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2936

QY 1699 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1758

DB 2937 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2996

QY 1759 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1818

DB 2997 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3056

QY 1819 CTG 1821

DB 3057 CTG 3059

RESULT 13	AB007934	14491 bp	mRNA	linear	PRI 28-AUG-2003
LOCUS	AB007934				
DEFINITION	Homo sapiens mRNA for KIAA0465 protein, partial cds.				
ACCESSION	AB007934				
VERSION	AB007934.3	GI:34328013			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Homo sapiens				
REFERENCE	1 Seki, N., Ohira, M., Nagase, T., Ishikawa, K., Miyajima, N., Nakajima, D., Nomura, N. and Ohara, O.				
AUTHORS	Characterization of cDNA clones in size-fractionated cDNA libraries from human brain				
TITLE	DNA Res. 4 (5), 345-349 (1997)				
JOURNAL	9811662				
MEDLINE	9455484				
PUBMED	2 (bases 1 to 14491)				
REFERENCE	Ohara, O.				
AUTHORS	Direct Submission				
TITLE	Submitted (08-OCT-1997) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba 292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp, Tel: +81-438-52-3913, Fax: +81-438-52-3914)				
JOURNAL	On Aug 27, 2003 this sequence version replaced gi:20521060.				
COMMENT	Location/Qualifiers				
FEATURES	1..14491				
SOURCE	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="1"				
	/clone="hg0112362"				
	/tissue_type="brain"				
	/clone_lib="pBluescriptII SK plus"				
	/note="This sequence was obtained by subcloning of the DNA fragments derived from three cDNA clones (1 - 3992 was derived from hb07943, 3993 - 8100 was derived from ffl10348 and 8101 - 14491 was derived from pf00419)-This sequence was replaced that of hg01123 cDNA as a representative cDNA sequence for KIAA0465."				
gene	1..14491				
	/gene="KIAA0465"				
	<1..13303				
CDS	/gene="KIAA0465"				
	/note="Start codon is not identified"				
	/codon_start=2				
	/product="KIAA0465 protein"				
	/protein_id="BAA32310.3"				
	/db_xref="GI:34328014"				
	/translation="KNTISVKANVDYROI EITICKNDECVLEPNSQRTKMKYISPTGN EAMVPSVCLIPPNKDAIEMASRVQSGOKWALMHOAHVNTKSLISMYLAKDLD VQTMNLEKRSAPGECDOI MNKLQAHYEDFDLSDSDVLFVADRLRLLEEVEACKA RQHLMKSMENEDKEETVAKMYISELKNIRLLEEYQVVKRISOIASRRDRDMQ DNALRIAEQHEHDEDLQOQLSDLDVASKSCDSFPLSSPSVPTLRSLNLTVEEMD HYYGLSTVYLNKLTVDIVYRSIODAEILVKGVEIRLSQEEVVLADLSALAEWSTLR HMLSDYKNSVSVSLDEIRAKKVAEKMRSRLTPERNLDIEYQKSGSLOERHVR IAOLEIROSELISIOEVLGDYRACHGTLIKWIEETTAQGBMKPGGAEDSRVLSBUS OQTALFABIEIRNOTKLDQCKQSQSTVVKDEILMTYKKAVEREQSGPGRKRLM SSSDALTOEFMDLTRRYTALVTLTTOHVYVISPALRLLEEKVEEKEQEHVEKYE LKHWGSTLARNOTGKATSEKSTDI EKALILEOVLSEELATRKQVEBATISQIF LAGHGLTSEKEKQISBOLNMLNKAYHLDLSDSANOLQOLQSLAHOTBOKTLQKQ NTCHOLEDLCSWVGQAEERALAGHGRRTQODLSALOKNOSDLKIDIDIONRATSEFAT YVVKDIEGFMEBKQTKLSPRELALAEKLEHQAKEYEALQEBTRVAKQELBEVVISAL QOETEKSAKAEILAENKKKI DALDLVNTVSGSGQLTNLPMEQLSGASLEKGLD TTDGVGVQAQPEPKLDKQCEMMKARHQEILSQONFILTQSOAQLDGHNLPTEE OOMLOQKLELKEYOYSTLAQSEALKEQVLODELQKQELDHKEEYVESSEKLE INHKGSSPETLPSLKROGSPSEDIYSHKGNLPYTIISGQKIDMENSFKXGKESE IGNIVKDKLKDATERYTLAKSKCTRGSHLNLIGYHFPONSADSLQMMQACENY EKLSDTVASDPRVLOEQALTTKQLOEBELAEHGVPEKQAKVARDIMELEGEPADPR HVOETDSTLISHRQSTLSYSLAERSLLOKRIASQSVQSLSESLQSTIEVEQNLEGK QVSSLSGVIIOEALATNMKLDKODIAROKSSLEATREMTVRFMETADSTAAVYLGCLA				

EVSRPEQLQOQEKESLKLPLPOAENFEHLSCGLQOQFMENKSHMLASGNPDODI
THFPOQIDELNTEEMDQDENTLDTLEHLYTFLSCGFLALDQCHODVQVLRKPTLQ
KTYFEREKDASCOBOLDENRFLVTRFOKMLKETSIPPTETSMKAEKLEKQIEHLK
SLIDDMASKGTVIERINCKQDLENIJMEITTPDQCGKDLNLTIEODMADVNYKYEKL
GGYTLEROSLQAIILNRMEEVKEANSVTLQMSREKVKSHDAMSPTKTEYVAKQA
ESNKAVLAELBQNSPKI QVYKALAGLVTYVNSOAEWKNKI QEBLNRWEATEVT
VARORLEESASHLACFOABESQLRPWMLMEKELMVGVLPLSIDPMVLAQOQVPM
LKEPEARQOEQLNEAOGILTGPDVLSYQVKELOS INQKVELTIDKNSKSS
QIDDAIVKSTOYQELLQDSEKRVAVGSLVQSALSTQPEAVQOOLESETRSLQ
QLDHYEKAQTCDELSYLGBOYLDKELKLFYALPLQGLIEDLAADRIRNLQAL
ASTQOQNFDELRTLDKQSOQANCTI SAKLERLOSQJENEFQSGLNQHSST
EYIVAGESELSVPEGEKRTLQNVLPKSHWELSKTDROSRLDCMOKAKY
QNHVEDLVWIDDCAKSELVTLDPVLBSLSLSKMLMEVERERBLLETLNSAA
DILINSSEADDEGIDDEKAGINMDAVTEILOAKTGSLEMTORLREOEFPKTEK
KVEGAKHOEITFDALGSGASKNKLEKLAQOEVOLAPROYDLNFPQGVEDPAD
GSDAQLQDAEVAQOEPLEVGORVNSGVMMENLTQEGCHRCVREMFQSLADLD
ELDNGAIGRTDSDLOSIEDVRLPLNKHVILKDIIEASEACRHLLEEGTDLILGL
KLEELANKQCGKLTERRGARQOULELTGRVEDFYRKGLKLNADATTALEEALQV
VGEVEEILNQQLADFPMQEQVDPLQMLQOVNGIGQILQSGADCDVQGLEHME
EINARVNTLNKVAORIAQJELALHCGFQALBPLLSMLADTELINQKPSAEY
KYVKAQIOEQKLORLIDRKRATVMDLQAEGRILQSAELADREKITTGLOEESWT
ELLSYAAQKQLEDILVAKOPHETABEISPLSVTEKKLANSBVGOTAKIQOQI
IYHKAINEBTVKKNVDOAIKNGQLKQTIGEEVLLQEBKLDGIKTRFADITVTS
KALRTLQOROLATKFOSTYEBELTGMRLREVEBELATVSDGQSPTEQI PQOQKRELK
KEVMEHRLVLDIVNEVSRLLELVPRAREBGDLVSDNBOYKLVSDITGQRVEDID
AAIORSQOQAADALAVAEKRLMALGPIRLQODQTTAKQVAKAFSIDIRBK
DAMDLEFSHRSIPTGCGEOKTVLOEKTESLIOQVEALISLNSERYALEROVAVN
OFWEYKEELSPWIEETRALIALPSAIDHEDLOQOQEMRORRESIAHKHIDKL
KIGPOLKELNPEGEWNEBKTRAKENMAQICIEYRQRLALDEAASQOSTOTTERDK
IEMPLEETENLSSRLRMPPLIAEVDKIRECISDNKSAVELEKQLPSEALRKSE
LIGRSGADKDLAAKBIQDKLOVWFEMFMDIARAEERIKELTDLVLEAEKVEYMAA
LLTTIKQTDIVHDESPGIDPSIIKQVYEALETTKEEDGHELEFRLIGADLIF
ACGETEKEPERSIDEMNNAMLNKTKMERELKEDANAQVQYDQTLQAMEMDYN
TVYKLTMPVQVQTDVNTYQDQNMKEPFVRYOQOIEEKLNOHSELMKATADTD
NDIIRPLETELHLMENGEKTAHROHKLBSGLMALQGOHALBELMSVLTPTBELD
ACRPISGDEKLEVELEAKHNLKNDVLAQVATEVYNKAGNELLESADDSLSKR
LEAMQCMESVLQKTEBERQOQLOSTLOQAGHVEFLEDTLETRENSQLSASPKPG
LBEPTAREODIMELYSOLKAEETYNQLDGRMLLSRDSGSSKTEQSVALLQEL
KHWVSSKKEERKSLEBLNATEPONSLOFIMVLTABOINLASPNSPILNVL
SOIEBHVKYANVNAHRODITLIDQGNOKRLSGQOVVILKNLVSVQSRREKYO
RSIEGRSLDDRKAKQFHEAKWKILDLBEASLSDLELISNDPDKIKQVSRQ
EFQTKLQGVQPYDITTRTGRALKERTLLPEBSOKLNDLFGAVRDKMDPCKGSLVQ
HKLLEBALFSGQFMDLQALVDMLYKVEVGLDEDPVHGDLDLVNMLAHKAYFQEL
GKRTGVQVLYKRSGRELLENSRDPVTWVGQLOEILSTRDYDCKLSVSKSRLQOLK
QAEVFRDPTVHMLLEMLSEAOITLRFGLPDLDTLALOSLIDHMKEMKVEEERDVN
SAVAMGEVILAVCHPDCTTTIKMTTIIIPAREEVLITNAKQKHLETLSELVANAE
LLEELAMVQMAETLLIQDQEPITQNDIVVALIAEHOTFMEBTRKQPDVDRVYXT
YXRKNIPEPTHAPFIEKRSRGRKSLSQPPMPMPLISQSEAKNPRINQJLSAKMOQVL
LALEROKLNDALDLLEBLEKFAFNDFDVMRYKRWMMNKKSRVMDFPRRRIDQDQD
KITROEFIDILASKEPPTKLEMTAVADLPDSGQGYIDYEFVLAHFNKQAVRETT
DADKTDEVTBROVACCKAKRPOVVOIGENKYRFPDSOOLRVLRILSTRVMMRVGGM
NALDEFLVANDCRAGRNTILREKFTILPEBASQGMTFRSRRGRSKRSRAASTR
SSSSASQSNHSCITMSPSPATPASGTVIPSSGSCLKRPTVPHSRSTSLADPTSSS
SPASGAKTRNADPKPSASRPGSRGSRGSAASRSRDPADFDLELTQSCASPTSE
SSAAGQGSRSRGLNKPSKIPTMSKTTTASPRTPGPKR"

ORIGIN

Query Match	32.3%;	Score 655.4;	DB 9;	Length 14491;
Best Local Similarity	62.9%;	Pred. No. 3.9e-148;		
Matches 1134;	Conservative	0;	Mismatches 645;	Indels 24;
			Gaps	7;
Qy	41	CCGCCCCCAGCCCTTGAATATGATCTCTAAGACGACGACGAGGAAACATTCGGCAACTG	100	
Db	8555	CCCTCTCCAGCCATTCATGATGACGCTCAGGCAACGACMAAGGAAATGAGGCAATT	8614	
Qy	101	CGTGAATTGATAGCTGACACAAAGCCTCATATGATTAAGATGAACAAATCGGCC-----A	156	
Db	8615	AAGGAATCATATTGCTGACACAAACCTCATATTGACAAACTATAAGATTAAGGCCCA	8674	
Qy	157	CAGTACTGATAGAGCCCTGGGGAAGGCTTTTCTATCCAAAGAGAAATGTGGCAGCCGAC	216	
Db	8675	CTTAAAGGAATTAAACCTTGAGGAGGGAATGTGGAAAGAAATAATCCAGAAAGCGAA	8734	

QY 217 ACCCTTACAGTCACATTAAGAGATGTCAAAAAGCGTGTGTCAGTCAGTATGAGCC 276
 DB 8735 AACATGATGCCCAATTAAGAGAGAGTGGCCAGCCAGCCCTGGCTCTGATTAAGCC 8794
 QY 277 ATTTCTCAATCACTC-----AGTTCCATGACAGATAGATGATCAGATCCTTGAAGC 327
 DB 8795 GTGTCGCCAGTCCACAGATTAACAGAGTTTCATGATTAATAATTGAGCCATGTGTTGAGACA 8854
 QY 328 CTGGAAGCCATCGTGGAAAGCTGTAGAGGAGCAACCCCTATCTCTGACAGAGTTGAGAG 387
 DB 8855 CTGGAAGATCTTCTCTCTGCTGCTGATGCGACACATGATCCCTGCTGAAGTATACAG 8914
 QY 388 ATCAGAGACAGATCAGTGAATAATAGATGTGTGATGACATGAGAAAGCTACAGCCG 447
 DB 8915 ATCAGAGAGTGCATCAGTGAACATTAAGAGTCCACCGTGGAGCTTGAAGAACTGACGCA 8974
 QY 448 TTGTATGAACTCTTAAACAGAGGGAGAGAAATGATGTCTAGATCTGGGGGAGCTGAT 507
 DB 8975 TCTTTTGGAGCCTTGAAGCGCGTGGAGAGAGCTTATTTGACATCTCAGGAGACAGAC 9034
 QY 508 AAAGACATATCTGCCAAAGCTGTACAGATTAAGCTTGACCAAAATGGTTTTCATTGGAG 567
 DB 9035 AAGATCTGCTGCAAAAAGAAATCCAGATTAATTTGATCAAAATGGTATTTCTTGGAG 9094
 QY 568 AACATACACACATGCTGGAGAGAGAGAGCCAAACTAGATGATGATGAGAGTACCA 627
 DB 9095 GACATCAAAAGCTCGGGCTGAAAGACAGAAATCAAAATTTCTGATGTCTTGAATTTAGCA 9154
 QY 628 GAAAGTTCTGGTGTGATCAGATGTCAATTTGATTAAGTCAATTAAGATCAAGATTTTC 687
 DB 9155 GAGAAATTTCTGGTATGACATGAGCAGCTCTCTGACACATCAAAAGACACCAAGATATT 9214
 QY 688 ATCCGGAGCTCGAAGATCTTGAATTTATCTTCAAGTAAAGAAACACAGCAAGAGCA 747
 DB 9215 GTCCATGATCTTGAAGAGCCAGGCAATGATCTTCCATCATCAAAACAGATTTGAAGCT 9274
 QY 748 GCAGAGACCATAGAGAGAAATAGATGATGATGATGATGATGATGATGATGATGATGAT 807
 DB 9275 GCTGAGATATTAAGAGAGACAGATGATGATGATGATGATGATGATGATGATGATGATG 9334
 QY 808 CTAGATCTGAATCTATTCGCGCATGTGGGAGCCTGTATTAACCATTTGTCAAGAGAGT 867
 DB 9335 CTTGAGAGAGATTTGATTTTTCCTGTGAGAAACTGAGAGCCGTGAAGTATGAGAGAGC 9394
 QY 868 ATACATGATTTAAATTCAGCGATGGATTTCTTAATTAATTAAGCTTGAAGAGCCGATGA 927
 DB 9395 ATTATGATGATGAATTAATGC-TTGGGAAACCTTAACCAAAACATGAGAAAGAGAGCTGAG 9453
 QY 928 CCAACTTGAG-----AGCATGCAAGTCCGCTGATGATGATGATGATGATGATGATGAT 981
 DB 9454 AAAACTTGAAGATGCTATGCAAGCTGCTGCAAGTATGAGACACTTTCAGGCTATGTT 9513
 QY 982 TTTCTGGATGATATTGCAAGTGTGTAATGCTTCATGTCTCAATTTGGAACAGCATC 1041
 DB 9514 TGACTGGGTATATAACTGTGATTAACCTGACCACTGCCCCCTGTGGCAGCTG -ACC 9572
 QY 1042 TCGAAACTGTCAAGACAGATTTGATGTCTGAGACAAATTTAAGTCTGAGCCCTATCAAC 1101
 DB 9573 TCAATTAAGTAAAGATCAGTTAAATGAAGAGAGTTCAAAGTGAAGATTTTACCAAC 9632
 QY 1102 AGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
 DB 9633 AGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9692
 QY 1162 NAGAGAGTGAACAACAACCTGTCTAGACCCAAATAGANCTGATTTGATAT -GGCTA 1220
 DB 9693 ATGAGAGGAGACAGAGACATTTATAGAGAACCACTGACAGAACTCAAAACACTCTGGAGAG 9752
 QY 1221 GCTTGGAGNAGAGAAATCATCAACAGACAGATTAATCTGAGAGGCTCTTATTTAGCCNTGG 1280
 DB 9753 ACCTGGGGAGAGAAATTTGCCCAACGACAGCAAAACTGAGAGGGGCTCTGTTGGCCCTTG 9812
 QY 1281 GTCAAGTTCCAAACATGCGCCTGTATGATGCTCTGCAAT -GCTACACACACGAGGCTTGC 1338

DB 9813 GTCAAGTCCAGCATGCGCTTGAAGAGAACTAATGATTTGGCTACTACATACCGAAGTGTGT 9872
 QY 1339 TAAGTACAGAAACCTGTTGAGAGAGACCTTAAGCCATTGAAATTTGAATCTTGCAGAC 1398
 DB 9873 TAGATGCTCAGAGACCAATTAAGTGAAGACCCAAAGATCATTTGAAGTGAAGCTCCCAAGC 9932
 QY 1399 ATCATGTGCTCCAAATAGATGATTTAGCCCATGCTCCAGTGAAGAGCGGTTAATTAAG 1458
 DB 9933 ACCATGCTCTTAATAAATGATGTTTGGCTCATCAAGCCACAGTGAACAGTCAACAAAG 9992
 QY 1459 CAGGAAATGATCTAATTTGAATCAATGATCAGAGAGAAAGCAAGCACTTTCAGAAACAGC 1518
 DB 9993 CTGCAATGAGACTCTTGAATTCAGTGTGAGATGATGATGATGATGATGATGATGATGATG 10052
 QY 1519 TAGAGTTTAAATCAACCTGSCAAAATGTTTGGAAAAACAAACAAAGAGCAGC 1578
 DB 10053 TGAAGCCATGAACCAATGCTGGAGATCAGTGTATACAGAAAAACAGAGAGAGGAGCAGC 10112
 QY 1579 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1638
 DB 10113 AGCTTCACTCACTCTCAGAGAGCCAGAGGCTTCCACAGTGAATTTGAAGATTTCTCT 10172
 QY 1639 AGTGGCTGACTGACACAGAGCGCTCATCTGTGATGATCAAAACGCTGGAGAGTTTACCGG 1698
 DB 10173 TGAAGATTTAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10232
 QY 1699 AAACAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1758
 DB 10233 AAACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10292
 QY 1759 AAGAAACATTAAGAGTCTGATGAGAGAAAGGCGCAGATGCTTGCAGAGAGAGAGAGAG 1818
 DB 10293 AAGAGACTTATTAATCACTACTTGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 10352
 QY 1819 CTG 1821
 DB 10353 CTG 10355

RESULT 14
 AB029290 17518 bp mRNA linear PRI 03-SEP-1999
 LOCUS
 DEFINITION Homo sapiens mRNA for actin binding protein ABP620, complete cds.
 ACCESSION AB029290.1 GI:5821433
 VERSION AB029290.1
 KEYWORDS actin binding protein ABP620.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS Matsuda,S. and Okuda,T.
 TITLE Actin binding protein ABP620
 JOURNAL Published Only in Database (1999)
 REFERENCE
 AUTHORS Matsuda,S. and Okuda,T.
 TITLE Direct Submission
 JOURNAL Submitted (22-JUN-1999) Satoru Matsuda, Nagoya University School of
 Medicine, Molecular Pathogenesis, 65 Tsurumai-cho Showa-ku, Nagoya
 466-8550, Japan (E-mail:smatsuda@bunru.med.nagoya-u.ac.jp,
 Tel:81-52-744-2463, Fax:81-52-744-2464)
 FEATURES
 source
 1. 17518
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HePG2"
 1. 17518
 /gene="abp620"
 52. 16344
 /gene="abp620"
 /codon_start=1

/product="actin binding protein ABP620"
/protein_id="BAA83821.1"
/translation="MSSDELTSESRCSERSCRSERYSRSGSLSPCPDITP
MNLPIHROKRRKSDSLDPABAAVVRVADDERDVOKRTPTVKNHGLMKVRHIND
YEDLRGHNLSLLEVLSGKLPREKGRNFRHLQVQIALDPLKORQKLVNRIND
ITDNGKLTGLIWTIILHFQISDIISGSGMSAKKELMTQRTVGYTGIKCN
FSGSGDKGMFNALHRRYRLDVMERVOQSRENLEQAFVARELAVTLADAD
DVSPDEKSVITVYSIYDAFPVPEGEIGISATVDSRWQEOSRVDSIIPMKOT
LMSDKTFPONVEALKALVNOYIHFKEETELAKEREKRIEELYKLEVEIGRIKL
PGYHNDVBEEMGKLIEMLEBEKSLPAVERLELILQIANKONGALNCPGKTLA
KNTLQDAHLEBSGVPVQCSVDYMTIQCEGILROLQDVLQRLRANTYQLEELFR
VMRLQELVTLRLCTNLIRKGFILSLVPEPSTLTTLHKAPLTKRNTSSVFR
KPTRAELDPSAEKAKMLASDLQNTYKMWQMLERAMENDPSEVLQETQSH
HTSVLEISSVSEARLYGKMSQNFHTSAETLQKLETOYCLAKETSFRMLQSH
HKFSRATAEILWLNKEBEELAVDSQNNMNSI SAKRNVSELTMLEEQRERLQ
DRAELSLNHPAKOTVEAVSAVOSLOMKNQCLCEVQHYKENTAYPOFSDAEL
ESFRLQDSIRKRTSCDNTSLRLLEDLQDSMDKEQLIQSKSVASLVGRKTIY
QKPRSFDHVKNTISVAVCDYRQIEIYICNDECVLEDNSQRTMKVISPENAM
VBSVCLIPPNKDAIEMASRVOSYQKWMALMHQLVNTKSLISNMYLRKDILVQ
WNLKRSASAPGCHQIMKQLQAHYEDFLQDSHDSVLFVADRLREBEVACAKAFQ
HMKSMENDEKEETVAKMYISELKNIRLREYRQVVRKIOSLASRDRQDMONA
LTAEOHNOEDLOLRSDLVASMKCDSTFHSBSSSVPTIRSLNLTLYEMDHY
GLSTVYLNKTYDVIYRSIQDAELVKGIEIKLQOEVLADLSALEHAWSTLRHL
SDVKDKNVPSVLDEBIAKAVAEQMSLTPERNIDLERYQKSGQLQERHVRITAO
LEIROBELSIOEVLGYRACGTLIKWLETTAQOEMKPGQABDSRVLSQLSQOT
ALFABIERNOTKLDQCKFSQSYTIYKQYELQMTYKAVESQKSPGRMRMSS
DALTOEFMDLRTRYALTALVTLTQHVAKIIGDALRLREBEKVEEKEHVEYKELIG
WSTLARNQKATSETSESTDIEKILIEQVLSBELTTKKQVEBAIKASOPIAK
HGHKSEKKEKQISEOLNMLNKAHYHLCGSAQLQDQSLKHQTEQTLQOQNTAK
HQQLQCSWVGQERALAGHGRITQOGLSALQKNSDLKLDQDIQNRATFAYV
KQIEGMBENQTKLSPRELTALEKLOAKQEOYALQOEBTRVAKQLEBAVSLQOB
TEKSKAKELAEKKKIDALLDWTISVSGSQNFLLNLMQMELOSGSLAEKALDQTD
GNGVQVAPKLDKCKEEMKARHOLLSQONFLIATQSAQELADQHGHNTPREOM
LOOKGELKEOYSTIASOEBALQVOTLODELQKLOHKEZESLSESEKLEMH
KGGSPPTPSILTKRQSGSEDTYSHKQDLPFTISGQVLMENSKFKGSESLGN
LVNDKLDATERYTLAHSKCTRGSHLNNLQGYHOFQNSALQAMQACANVEXL
LBDTASDPGVLOEATTKQOELAEHQVPEKQKQALDIMEIEGEPADGRHVO
ETTDSILSHFOSIYSLAERSLSLOKAIQOSQVDSLSLOQISEVQENEDRVS
SLSGVIOEALATNMKIKODIAROKSLETRMTYRPMETDASTTAAQDILAVS
ORPEOLCLOOEKESIKLLPOAEWPEHSGLOPMMKSMMLASNOPODITHP
FOQIOELNMEQOENLDTLEHLYTSLSCGFALDCOHQDQVNLARQDPTBELQTV
KREKQASQEOBQDFRLVTFQKWLKTEGSIPTPTSSAKLEKQIEHLKSL
DMWASKGLVEEINRYGSIENLIMEITAPDSQKTSILPSSGSGVNGVGHITKD
LLEIQCMDMDVNLKELGCVLHERQESLOALINMBEYHKNESVLYQWLESEELK
SDMSSPTKETVKAQASNKAFIAELINSONPKIOXKBALAGLIVTAPNSQEAAM
KKIOEILNMBEATETVVARORLESASHACPOASSQOLPMLEKELMGVIGP
LSIDPMMLAQKQOVPMKKEFBARQOHLNEAQGLITGEGDVSLSQVQKLO
SINQWVELIDKUNRSQIDQALVASTYQBELLOQSEKVRVAGRLSVQSAISTOP
EAVQOOLEBTSEIRSDLEQLDHEVKEAQTLCDLSTYIGBOYLKBLKRLTVAPL
OGLEVDLAADRINRLQALASTQOFOQMPELTLDLDKQSOQAKNCPISAKLERLOSQ
LOENBEFQSLNQHSGSEYIVAEGLSLSVPGKEKTLQNOVLKLNHMEISKK
TADROSRLKDCOKQKQYMHEDLYPMIEDCKASBLRPTLDPEVLESSLRBAM
LNEVXRSLBLIELNSADILINSSPADBDGIRDEGAGINQNDADTEBLQAKTSLB
EMTQRLREFQESFKNLEKVEGAKQOLIFDALGSAQCNKMLKELRAQOEVQALBP
QVUYLNFQGLVEPADPDSQOLHQAQVAVQOEFLEVKORNSQEMMEKLEBEG
QCHCRVEMFSQALADDELDMGAIIGRPTDSLOSQIEPVRLFKINHTKLILKSE
ACRHNMLEEGTLDLGLKRELEALNKQCKLTBRGKABOBLIELGLVEGPPYRLKX
GLNDATTAAEBAALQWYVTEVEILNQOLADPKMQKQOVPLQKLOQVNGGGL
IOSAGDCQVQIEHDMETINAMNTLNKVRQIRLQLOEBALHQAQEGRDALPELW
LADTEBLIANQKRPESAVYVKAQIQOBLQRLDGRVADYDMQAOBGRALQASLE
ADREKLTQGLSESRWTELSKAAARQOLELDIVLTAQFQHSISLSPBQVTEK
LANSBPVGTQAKIQOQIRHKLAEEDINHAIDYQAVKIOSQISLSPBQVTEK
EKIDSLOARYSEIODRCRKAALDQALDNARLFGEDVEVEMTLAEVDKISVAVK
DFKODIAROHADHLALNEIYVRKKNDVQAIKNGQALLKQTTGEBVLLIOEKLDSIK
TKYADITTSKALRTLEQARLATFQSTYBELTGMLEVEBELTASGQSPTEGQI
POQVQKELKXVMEHRYVLDTVNEVSALBELVPMRABGIDKLVSPANQOYLQV
DTIQGVDEIDAIQORSQOYEQADAEILMVAETKRLAALGPIRLQEOQTAQKLV
KAFSIDIIRHXSMDLFSHRSIEFGTCEBQKTVQOETSESILQOYEAISLNSRY
ALERAQVAVNOFMEYTELELSPMIEETRLIAQOLBPBALDHQLOQSEMOLEBSI
ABHKPHIDKLKIGPOLKEINBEGEMVEEKQKAKENMTAQIKEBTRQALALDEAVS
QSTQITFERPDKIEPMLBTLENSSRIRMPPLPAEVDKIRECISDNKSAVLEEKQOP
SEALKRREBELIGSOGADKLAKEIDQKLDQWVFWEDIKARAEEREIKFLDVLE

ORIGIN									
Query Match	32.3%	Score 655.4	DB 9	Length 17518					
Best Local Similarity	62.9%	Pred. No. 3.9e-148							
Matches 1154	Conservative 0	Mismatches 645	Indels 24	Gaps 7					
QY 41	CCCGCCCCACCCCTGGAATATGATCTCTAAGGACGACAGAGAAAGATCGGCAACTG	LAKEFWYDAAALLTTIKDTQDIDVHLESPIIDPSIIKQOVEAETIKEBTDLHEELE							
DB 11596	CCCTGCCAGCATTTGATGATGAGACGCTCAGGACGACAAAGAGAAATGAGGCAATTA	FLIRIGADLIPACGETEKREBARKSIDENAMNENIKTKWEKLEDMQOAVOYOD							
QY 101	CCTGAGTTGATGCTGGAACAAAGCTCTCATATGATTAAGATGAAGAAACAACTGGCC	TLQAMGLDMLDNTVILKCTWPPVGTNDATNKQDLENKKEKREVEYQOQIEMEKLNQGE							
DB 11596	AGGGAATCTATTTGCTGMAACAAACCTCATATTGACAAACTAAGATAGAGCCACAA	IMLKATDSTEDSDIIREPLTKHMEINCEKILARHQLBEAALLAGFORALEELM							
QY 157	CAGTCTGAATAGAGCCCTGGGAGAGGCTTTCTATCCAGAGAAAGATATGGCAGCCAG	SWLTTREBLDORPISQDEPKIYBVLAKHNTLVADYLAHQATVEITVYKAGNELSS							
DB 11716	CTAAAGAAATTAACCTCTAGAGAAAGGGAATGCTGGAGAAATAATCCAGAAAGAGAA	ACDDASLSRLLEANNQCSVEYLQITEEERQOLQNDYLAHQATVEITVYKAGNELSS							
QY 217	ACCTTTACAGTCACATTAAGAAGATGTCAAAAAGCGTCTGTGGCACTGATGAAGCC	SOLSASKPPGGLPETAREQDLMELYSQLKAEETVNLQDLKGRMLMLSDSGSGS							
DB 11776	ACATGATAGCCCAATAAAGAGAGAGTGCCGACGAGCCCTGGCTGTGATGAAGCC	KTEQSAKPLPGGGLPETAREQDLMELYSQLKAEETVNLQDLKGRMLMLSDSGSGS							
QY 277	ATTTTCATCAATCAATC-----AGTTCCATGACAGATGATGATGATCTTGTGAGAC	SPESILNVLNQLIEBHKVFANVANAHRQIILEQDTQGLKFLSOKOVVILKNILVA							
DB 11836	GTGTCCCACTCACAAGATTACAGAGTTTCATGATTAATAATTGAGCTTATGTGGAGACA	SVOSREKVOOSIERGSLRFLANDBARQAFHAKMTDMLDASHSLHLESTISNDP							
QY 328	CTGGAACGATGCTGGAAGCGTGTGAGGACGACACCCTATCTCTGACAGAGTTGGAAG	DKIKIQLSHKRFQUTLGQKQVYVOTTTITGALKKEXLILPDSQGLNDFLEGVRKW							
DB 11896	CTGGAAGATCTTTCTCTGCTGCTGCTATGCCACCACTGATCCCTGCTGAAGTGAAGA	DYVCGKSVRQKLEBALFSGQFMDALQVDMYKVPQALBDDPVAGDVLVNL							
QY 388	ATCAAGAAACAGATCAGTGAATAATGAATGTGTCAGTAGACATGAGAAAGCTACGCCG	MOAHKVPQELKRTGTGYVQLKRSQELIENSRDRTYVWKGQLOEISTMDVYCKSLV							
DB 11956	ATCAGAGATGATCAGTACATTAAGATGCGCACGTTGAGACTGAAGAAATCGACAGCA	SKQSRLEQLAKQAEFVRLVHMLLWLSAEQOTLSPFGALPDTLEALQSLIDITHEFM							
QY 448	TTGTATGAACCTCTTAACAGAGGGGAGAGAAATGATTGCTGATCTGGGGGAACTGAT	KYBEERVDVNSAVVAMGEVILAVCHPDCTITTKHMTTIRAFEEVLTWAKHOQOLE							
DB 12016	TCCTTTGAGGCTTGAAGCGCTGAGAGGAGGCTTATTTGACGATCTCAGGAGACAGAC	TLASLVNAAELLEHMLIOWAETTLIORDPEIIPONIDRVKALIAHOTHMEETR							
QY 508	AAAGCATATCTGGCAAAAGCTGTGAGATTAAGCTTGAACCAATGTTTCAATTTGGGAG	KQPDVDRVYTKYTKRKRIETPAPRIEKSRSGRKSLISQTPPPMPLISGSEAKNPIIN							
DB 12076	AAGGATCTGGCTGCAAAAGAAATCCAGATTAATTGATCAATAGGTATCTTCTGGGAG	QUSAKQOVMLIALERQRKLDNDRLEBLKEFANFDFVMWRKTYRMNHHKSRVMD							
QY 568	AAACATACACACTGCTGGAAGAGAGGAGCAACTACTGATGTGATGAGAGTACAGCA	FPFRJDKQDDGKITRQEPIDGILASKFPPTKLEMTAVADIIPRDGQYIDYVEPVAL							
DB 12136	GACATCAAAAGCTCGGGCTGGAAGAGAGAAATCAAAATTTCTTGATGCTTGAATTAGCA	HNKDAYRPTTDADKIEDEVTROVAKQCKAKFPQOISENRYRFGDSQQLRLVRLR							
QY 628	GAAAAGTTCTGTGTGATCATCATGTCATTTGATGATTTACATTAAGATCTCAAGATTTC	STVYRVGGGMMALDBFLVYNDPCRARGRTYELREKFTLPPGASQGMTPPSPGRSR							
DB 12196	GAGAACTTCTGTATGACATGCACTCTCTCGACACACCATCAAGAGACCCAGGATATT	KRSRBAASPTBSSSASQSNHSCIMPSSPATPAGCTYIPSSGKLRPTPTPHSR							
		TSIAGDTSNSSPATGATKNRADPRKTSASRSGSRGASRAGSRASRSGSDASDPDL							
		ETQASCDTSESSAAGQGNRSRGLNPKSLIPMKRTTTPASPTTPPPRR"							

Oy 688 ATCCGAGACCTGGAAGATCTGSAATTGATCTTCAGTAGTAAACAAGACAAGACA 747
 Db 12256 GTCCATGCTTGGAAAGCCCGAGCATGATCTTTCATCATCAACCAAGGTTGAACCT 12315
 Oy 748 GCAGAGACCATAGAGGAAGAAATAGATGAGCTACAGAGAGAGCTGATATATTATTAAC 807.
 Db 12316 GCTGAGACTATTAAGAGAGAGACAGATGCTGCACTGAAGAGCTGAGATTATTCGGATC 12375
 Oy 808 CTAGGTTGAACTCATTCGCGCATGTGGGAGCCTGTATTAACCCATTGTCAAGAAAGT 867
 Db 12376 CTGGAGACGATTTGATTTTGGCCGTGTGAGAAACTGAGAAAGCCGTAAGTAAAGAGAGC 12435
 Oy 868 ATACATAGTTAAATTCAGCGATGCGGATCTCTAAATTAAGCTTGAAGACCCGATGA 927
 Db 12436 ATTGATGATGATTAATATGC-TTGGAGAACTTAAACAAACATGGAAGAGAGGCTAGA 12494
 Oy 928 CCAACTTGAG-----AGCAATGACAGCTGCCGTCAGTACAGATGAGCTGACAGCGGTATT 981
 Db 12495 AAACTTGAGGATGCTATGCAAGCTGCTGCAAGTATCAGACACTCTTCAGGCTATGTT 12554
 Oy 982 TTTCTGGTAGATATTGCAAGGTGTAAGTTCGCTTCAATGCTCCAAATGGAACAGATC 1041
 Db 12555 TGACTGGCTAGATTAACCTGTGATTTAACTGTCACATGCCCTGTTGGCACTG-ACC 12613
 Oy 1042 TCGAACTGTCAACAGACAGATTTGATGTGCTAGAGCAATTTAAGCTGAGCCCTATCAAC 1101
 Db 12614 TCAATACGTTAAGATCAGTTAAATGAAGAGAGCTTCAAGTAAAGTTTACCAAC 12673
 Oy 1102 AGCAGATAGAGTGTGATCAGCTGATCATCAGCAGACCCCTTGTGTAAGAAAGTAAAG 1161
 Db 12674 AGCAATTTGATGATGAGAGAGCTTATATCAGAGGAGTATGATGTTAAAGAAAGCTATCTG 12733
 Oy 1162 NAGAGAGTGAACAAACACTGTTCTAGACCCAAATGAGCTGATATTGAT-AGCTA 1220
 Db 12734 ATGAGACGAGAGAGACATTTATACGAGAACCACTGACAGAACTCAACACCTTGGAGAG 12793
 Oy 1221 GCCGTGAGNAGAGATCATCAACAGACAGCTAAATCTGAGAGGTCTCTATTAGCCTG 1280
 Db 12794 ACCGTGGTGAAGAAATTTGCCCAACGACAGCAACAACTGAAAGGGGCTCTGTTGGCCCTTG 12853
 Oy 1281 GTCAATTCACATGACCTGTATGAGCTCCTGCAAT--GCTGACACACACGAGGCTTGC 1338
 Db 12854 GTCAATTCACATGACCTGTATGAGAGAACTAATGATGCTGATCTATCCGAAGATTGT 12913
 Oy 1339 TAACTGAGACGAAACCTGTTGAGAGAGACCTTAAAGCCATTGAATTGAATCTGCAAGC 1398
 Db 12914 TAGATGCTCAGAGACCAATTAAGTGAGAGACCAAAAGTCAATTGAAGTTCGCAAGC 12973
 Oy 1399 ATCATGTGCTCCAAATATGATATTAAGCCATCAGTCCACAGTGAAGCCGTTAATTAAG 1458
 Db 12974 ACCATGTCTTAAATATGATATTTTGGCTCATCAAGCCACAGTGAACCAATGCAACAAAG 13033
 Oy 1459 CAGAAATATGATCTAATTAATCAAGTGAAGAGAAAGAGACCAACTTTCAGAAACAGC 1518
 Db 13034 CTGGCAATGAGCTTCTGAAATCCAGTGTGAGAGATGATGCCAGAGCTTAAAGAGCCGTT 13093
 Oy 1519 TAGAGGTTTTAAATCAACGCTGGCAAAATGTTTTGAAAAAACAAGACAAAGAGAGCAGC 1578
 Db 13094 TGAAGAGCATGAACCAATGCTGGAGAGTCAAGTGTAAACAGAAAAACAGAGAGAGGAGCAGC 13153
 Oy 1579 AGCTGAGTGTGCTTGGCCGAGGCAAGAGGTTTCAATGCGGAATTTGAGATTTCAGC 1638
 Db 13154 AGCTTCACTCAACTCTGAGAGAGGCGCCAGGCTTCCAGAGGAATTTAAATTTCCCTCT 13213
 Oy 1639 AGTGGCTGACTGAACGAGAGGCTCATGTTGGCATCTTAAACCGCTGGAGAGATTTCACCG 1698
 Db 13214 TGAAGCTTACTAGATGAGAGAGGCTTTCTGCACTTAAGCCACAGAGAGAGATTCTCTG 13273
 Oy 1699 AAACAGCCAAAGAGAGCTTAAATGCTCATATGAGAAAGTGTGCTGCTTTGAAGCTTAAAG 1758
 Db 13274 AAATGCTAGGGAACAGCTGATACACATATGAACTCTATTTCCAGCTGAAAGCAGAGG 13333

Oy 1759 AAGAACTATATAGAGTCTGATGCGAAGAGCCAGCAGATCTGTCAGAGATGCCCAAT 1818
 Db 13334 AAGAGCTTATTAACACTACTTGAACAGGACAGACTCATCTTAAAGCCGTAGACACT 13393
 Oy 1819 CTG 1821
 Db 13394 CTG 13396

RESULT 15

AF141968 17734 bp mRNA linear PRI 08-NOV-1999
 LOCUS Homo sapiens trabeculin-alpha mRNA, complete cds.
 ACCESSION AF141968
 VERSION AF141968.1 GI:6273777
 KEYWORDS
 SOURCE

ORGANISM

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

1 (bases 1 to 17734)
 Sun, Y., Zhang, J., Knaef, S.K., Anciair, D., Chang, M.S., Liu, Y.,
 Sutherland, R., Salgia, R., Griffin, J.D., Ferland, L.H. and Chen, J.B.
 Molecular cloning and characterization of human trabeculin-alpha, a
 giant protein defining a new family of actin-binding proteins
 J. Biol. Chem. 274 (47), 33522-33530 (1999)

TITLE

JOURNAL MEDLINE
 PUBMED
 10559237
 2 (bases 1 to 17734)
 Sun, Y., Zhang, J., Anciair, D. and Chen, L.B.
 Direct Submission
 Submitted (10-APR-1999) Cancer Biology, Dana-Farber Cancer
 Institute, 44 Binney, Boston, MA 02115, USA
 Location/Qualifiers

FEATURES

source

CDS

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="p33-p34.2"
 965..17086
 /note="cytoskeletal giant protein; actin-binding protein"
 /codon_start=1
 /product="trabeculin-alpha"
 /protein_id="AA06360.1"
 /db_xref="GI:6273778"
 /translation="MPPVLMAGIPGRVDSLOLPPEPKOCTSSASHVAAYADRDREY
 OKTPTKVMKHLKVKHINDIYEDLRDGHNIISLEYSGLIPREKGRMRHRLQ
 NVOALDPLKORVKLVNIRNDITDNPGLTGLITVTLIHPDISIYISGESGMS
 AKELULMTKQVTAIGYICINTESSCSWGSKFNALHHRPDLVMEKQISNRE
 NLQDAFEVSRILGTVRLDAEDVDVDPDKSVITVVSIVDAPKVEGEGISATE
 VDSRWQDYQVRVSLIPWIKQHTILMSDKTFPPQVPLKALYNQYHFKETELAKSR
 EKRIBELLYLLEWVIEFGRIKLPOGHNPDEBEMKLIEMLEERSKRPAYERLE
 LLIQIANKINGNLNCEBKLTLNKTLOAANLHESGQAPQCSDDVIMYIOEGGLIR
 OLVDLIQIIDENYVQLEELAFPMVPLQDILVLRLECTNLYRKGHTSLEIVPSLT
 TTHLRKAPLPTKATHSSSTISWFRKPRMTRALVAISSDEBGNLRFVELSWBEMQ
 KLEBAWGNLPSVELQLETOQHILHSVEBELSKSAPEGHQIMKYLQAHVEFLDYSRS
 KLITQYCKLKETESFRMRHLQSLHKFSPRTALEILIMNEEBEELADWDSNNISA
 KRVYFSELMELEEKQVPRSLQDTALLLENHPAQVTEAASAAQSOLOMKNQIC
 LYEQHYKENTAVAFQPSDARELESFLRNIDQSIKRYSCDHNTLSRLDLQDSMD
 EKQILOKSSVSIVGRSKTIYQIKRSRDYHLKTIYKAVQDVQOIEITCKNE
 CVLEBNSQRTKRVKISFTGNEMAVPSVCFILPPNKDAIEMASKEVDSYQVMAHIO
 LHVNTKSLISWNYLRKLDLVQTNLEKLSAAGECHQIMKYLQAHVEFLDYSRS
 LVPSVADRLEEBEVEACARFQHLMSMENEBKEETVAVKQYISELKNILRLREYEO
 RVVRIQSLASRTDRDAQODNLRILQEOHTOEDLOQLSDDAVSMKDSFLHQP
 SSSVPTLRSELNLVEMKMDHVGSLTVYLNKLTQVIVIRSIQDAELLVKGYEIKS
 OEVYVLADISALFAHSTLRHMSDVYDKSVSVSVDDEIARAKVAEOMSRILTPRN
 LDERRVQKSOLOERHVRVIAQLEIROSLESTQVLOVQCFRACHGTLIKVETTLQ
 QEMKPGQADSRLSELSQQTALFEIIRNDTKLDQCGKFSQOYSTIVKVELTQ
 TYKAFVSQOKSPKRRRRLSSDAIIOEPMDIRTYALVTLTQHVKQYISDLRL
 EEBKAVIEEKEQKVEKVELLQSVTLANTQOKATSEKSTEDIEKALIEQOVS
 EELTTKEQVSAIKTSQIFLAGHGHLSKEKKQISEQNLAKVAHYDLCDSANOL
 QQLSQALHOTBQKITLOKQNTCHQQLDLELCSWQGBRRLAHGQRTTQODLSALOK

NSGDLKODDIGNRATSFATVVKIDIEGFMEENQTKLSPRELTALREKHLQAOEYEA
 LOEETRYAOKLEBEAVTSALQOETSKAKAELEANKKIDALIDWVTSSGGGLT
 TNLPGMELOASKLEKALDITTDGYMGNAAPKLDIOCMGMAHOELLSSQGNPLT
 ATOSAOFLDHOHNLTPREOMLLOOKLGLKEOYSTLSAOSAEKONOTLODELOK
 FLODHKEFESMLERSEKELEENMKGGSPETLTERLYTALSKCTRLGSHLNLMLGYH
 SGRKVDIMENSFKGKPSSTGNLVKDKDATERLYTALSKCTRLGSHLNLMLGYH
 OFONSADSLQAMQACEANVEKLSLDVADPGVLOELATYLOQLOEHLAEVPAEK
 LOKVARDIMEIBEGEPADHRHVOETDLSLHFSQSLSYLAESSLSLOKLAOSQVQ
 ESLESLLOSIGEYONLKGQVSLSSGVLQKLAATMKKODIAROSKEHLETRAV
 TREMETADSTTAALQOKLAESVORPELOLOOEXKSSIKTLIDPOEMENHSGLO
 OFENKSRMLASGNODODTYTHFPOLOEHLMEHQOENLDTLEHLYTESCGRAL
 DLCOHODRVONKRDFTTELQITVEKSKDSSCOEQLDERKLVTFOKLKEBESI
 PPTESVSANKELEKQIEHLKSLIDMAKSGTIVEEINCKSTSEJIMETIADSOEK
 TDLTELOCMDSVNLKYEKLGVLHEKESLOALILNMEVEKANSVLOWLSKEEV
 LKSDMAESPTEKTEVYAQAESNKAPLALEONSPIKYEKELAGLVLTVPSOESE
 NMKIOEELSREBERATEVTVAROSLEASHLACQAOASIIILGPRDVSISTOVKE
 GPILDPNMILNAKQOVOPMLKEPEARROHEBOLANQOIIILGPRDVSISTOVKE
 LOSINQKWEZLTDKLSRSSOIQALIVKSTQYQELLODSEKRAVAGORLSVOSLIST
 OPEAVKQOLEETSEIRSDLEQIDHEVAEATLCELSVLIGBYLIDELKRLLETVAL
 PLOGLDILADRIINRLOALASTQOFQOMDELTWLDKQSOQAKRCPTLSLEBLQ
 SOLOENBEFQKSLNHOSSGYEYVIAEGESLISVPPGEKRTLOLNOVLKXNHEIS
 KKTADROSRLKCDMOKAKOYOMHVEDLVPIEDCKXMSLRYTLDPVOLESLSLASK
 AMUNEVKRRSLELILNSADLILNSSEADDEDTIRBKAGINOMDAVTELOAKTGS
 LBEEMTORLKEFOESFKNTEKVEGAGHOLETFLDALSQACSNLEKRLAQOEVLQAL
 EPOVDYLRNFTQGLVEDADPGSDASQLHQAEVAQOEFLVEKQVNSGCVMMENKLEIG
 IGOHCRVEMFSQLADLDDELDMGAIIGRDTLSQSLIEDVLPFKLILVLDLDEA
 SEAECHMLEEGETLIDLGLKRELEALNCKGGLTERGKROGLOELITLGRVDFPRK
 LKGLADTTAAEEBALOMVYGEVEEILNOULAPFMPOEYOVDLOMLQVNGGCO
 GLIOSAGKODNOGLDHEMBEINARWNTINKAKQRIAQLOEHLHCKGRDOLAEPL
 SHLADTLEBLANOKPSPAEKTVKAOIQOEKLOQLRLDDKATVMDLQAEKQIEABPL
 ELADREKRTGOELESLESRWTELLSKAAAREKQEDLILVAKOHEHAEPSISDSTE
 KKLANSBPVQTOAKIOOQIRHKLAEEDENATDVHQVKIGOSLSLETSABEGV
 LSEKISLQARSEIODRCCAKAALDOALNSARLFGEDVEVLMALAEDEKLSVFP
 VOFKODVLRHROADHLANEELVNRKNKWAQIKNGOALIKOTTEVEELIOEKIDG
 IKRRVADITYTSKRLPTLEOARQLATKOSTYBEITGMIREBEELANSQSGSPGE
 QIEPOFOOKELEKKEVMEHRLVDIVYENSRALBELVPMKARGDLKLSVDAEOYKL
 VSDITQRAVEDIDAIQORSOQYQADAEALVAETKRLMALGPIRLDQOTTAQLO
 VQAFSIDILRHKDSNDELFSHRESEIFGTCGEOKVLEDEKTESLLOQVLELILNS
 RYARLERAOVLVNOFWETYEELSPMEETELIAQPSPLIDHEOROOOEOROLSE
 SLAEHKPIEDKLKIGPOLKEALPBESEMEKYOKANENYAOIKREVAOENLALDEA
 VSGSOTITEPHDKIEPMLTEMLSESLRLEMPRIIPABVDKIRECISDKNSATYLEKL
 OSEFELKRGEBELIRSGADKDLAKELIQDLDMVFFMEIKARAEREIKFLIDV
 LELAEKFWYDMALTTIKOTODIVHDSLPGIDPSIRIKQOYVLEKIDAEEDLAEH
 LEBIRILGADLIFACETEKEPEVRSIDENMANMENINTKTRLEKLEIDMAQAOY
 ODTLOAMPMLDNTVILKCTMPFVGTDLNVKQOLMEKEFEVYOOOIMENKANO
 GELMKKATDETDROIIBRELTIELKELMNLGEBKIAHROHKEGALLAQOPHAEER
 LMSKLTHTHEILDQAPRISGDPKIVIEVELAKHVLKNDVLADHATVETYNKAGNELLE
 SAKGDDASSLRSLFLEMANOCWESVLQKTEREQLOSTLQOAGFSELEIDFLILTR
 MESQLSASKPTGLPETARQOLDTHMELYSOLKAKETVNDLDRKMLLSRDLDSGS
 GSKTEOSVALLBQMKHVSIMKERSKLEALNLATEFONSLOEFTINMLTLAEOSILN
 IASPSLILNLTQOIEEHKVPANEVNAHDOIIELDQGNOLKFLSODQOVLILKUL
 LVSVOERMEKVVORSIERGRSLDADARKKQOHEAWKILIDMEDESHLSLSELSIN
 DPDKILOLSKHEPOKTIAGKQPVDTTIRGRALKERTLLEPDKIDNFIAGEVD
 KNDTVCKGSEVEROHLKEBALFSGQFMDLQALVMDLYVEPELADQVPHGDLIDM
 NIMDAKRVQKEIGKRTGVQVLKRSGERLEIENSBDRTVWKOLOLSTETMDYCKL
 SVSKOERLOALQOAEVFRDVTMLLEMESEAOLEFRALPDTALSOELSDITHEEF
 MKKVEKRVNVSAAVMGEVILAVCHPDCTTIKHWITIIIRAFEEVLVMAQHOEMT
 ETALSELVANAELIEELLAMOWAETTLIORDOEPIDONIIDIRAFEEVLVMAQHOEMT
 ROPQVDVRYTKYKKNIEPTTHAPRIEKRSGGRKLSQTPPEMILSOSEKNEPI
 NQLSARWQVMLIALERKORLNDALDRLELEFANFDVWVKMTRMNMKKSVM
 DFFRRLDKQDGIKTROEFLIDGLASKPPTKLEMTAVADIPRDQDVIDYVEFVA
 LHPNDVAYRPTTDADKIDEDTVOVQAQCAKRFQYQIGENYRGSDSQOELVIL
 RSTVAVRGGGMMALDEFLVNDPCBARATNIELEKEFLPEYKAGGNTLPPSRERR
 SKXSSAASPTRSSSSASOSNHSCTSPSPATPAGSTVIBSSGSKLGRPTPTPSS
 RSLAGDTSNSSSPATGAKTRADPKKASAPRGSAGSAGSRRASRRGSDSDPDL
 LETOSACDSTBESSAAGGQNSRRGLNKSKLPTMSKTTTASPRTPGPKR"

ORIGIN

Query Match. 32.3%; Score 655.4; DB 9; Length 17734;
 Best Local Similarity 62.9%; Pred. No. 3.9e-148;
 Matches 1134; Conservative 0; Mismatches 645; Indels 24; Gaps 7;

QY 41 CCCGCCCGCCCTTGATATGATACTTAAGCAGACAGAGAGAACATCGCAACTG 100

Db 12341 CCTCTCCAGCAGCTTATGATATGAGCAGCTCAGGCACCAACAGAGAAATGAGCAATT 12400
 QY 101 CGTGAATTGATAGCTGAACAACAAGCTCATATGATTAAGATGAACAAACTGCC----A 156
 Db 12401 AGGGAAATCTAATGCTTAACAACAACCTCATATTTGACAAACTATGAAGATAGCCCA 12460
 QY 157 CAGTATCGATAGAGCCCTGGGAGAGGCTTTTCTATCCAGAGAGATATGTGGAGCGAC 216
 Db 12461 CTAAAGAAATTAAACCTTGAAGAGAGGAGAAATGTGGAAGAAATTAACAGAAAGCA 12520
 QY 217 ACCCTTTACAGTCACATTAAGAGATGTCAAAAAGCGTGTGCACTGATGAAGCC 276
 Db 12521 AACATGATAGCCCAATTAAGAGAGAGTGGCCGACGAGCCCTGCTGTGATGAAGCC 12580
 QY 277 ATTTCTCAATCAATC-----AGTTCCATGACAAATGATAGATCCCTTGAGAGC 327
 Db 12581 GTGTCCCAAGTCCACAGATTAACAGAGTTTCATGATTAATTAAGGCTTATGTGGAGAC 12640
 QY 328 CTGAAGCAGATCGTGAAGCGTCTGAGGAGCCACCTTATCTCTGAGAGTTGAAG 387
 Db 12641 CTGAAGATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12700
 QY 388 ATCAAGAAACAGTCAAGTGAATAAAGATGTGTCACTAGTACATGAGAAAGCTACAG 447
 Db 12701 ATCAGAGATGATCAGTGAATAAAGATGATGATGATGATGATGATGATGATGATG 12760
 QY 448 TTGTATGAATCTTTAAACAGAGGAGAGAGAAATGATGCTATGATCTGGGGAGCTG 507
 Db 12761 TCTTTTGAAGCCTTGAAGGCGGTGAGAGGAGCTTATGAGAGCTTCAAGGAGAGAC 12820
 QY 508 AAAGACATATCTGCAAGCTGTTGAGATTAAGCTTGAACCAATGTTTCAATTGGGAG 567
 Db 12821 AAGATCTGCTGCTGCAAAAGAAATCCAGGATTAATGATCAATGATGATGATGATGAT 12880
 QY 568 AACATACACACTGTGTGAAGAGAGAGAGCAAACTACGTGATGTGATGAGCTACGA 627
 Db 12881 GACATCAAGCTCGGCTGAAGAGAGAGAGCAAACTTCTTGATGATGATGATGATGAT 12940
 QY 628 GAAAAGTTTGTGTGATATCATATGATGATGATGATGATGATGATGATGATGATG 687
 Db 12941 GAGAAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13000
 QY 688 ATCCGGGAGCTGGAAGATCTGGAATGATCTGCTGATGATGATGATGATGATGATGAT 747
 Db 13001 GTTCATGATCTGGAAGAGCCAGGATGATCTTTCATATATCAACAAACAGTTGAAGCT 13060
 QY 748 GAGAGACATTAAGGAGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 807
 Db 13061 GGTGAGACTATTAAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 13120
 QY 808 CTAGGTTCTGAATCTATTCGGGATGTGGGAGACCTGATTAACCCATTGTCAAGAGAT 867
 Db 13121 CTTGAGAGCATTTGATTTTTCCTGTGAGAGAACTGAGAACCTGMACTGAGAGAGAC 13180
 QY 868 ATACATGATTAATTAATCAGCGATGAGATCTCTAATAAAGTTGGAAGACCGGATTA 927
 Db 13181 ATTTGATGATTAATTAATTC-TTGGAGAACTTTAAACAAACATGGAAGAGAGGTTA 13239
 QY 928 CCAACTTGA-----AGCAATGCACTGCGCTCAGTACAGATGACCTGACGCGGTATT 981
 Db 13240 AAAACTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13299
 QY 982 TTTCTGGTGAATATTGACAGGTGATGATGATGATGATGATGATGATGATGATGAT 1041
 Db 13300 TCACTGGCTGAATTAACATGATTAATTAATCTGACCAAGGCCCTGTGTGGACATG-ACC 13358
 QY 1042 TCGAACTGTCAAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1101
 Db 13359 TCAATGATTAAGATTCAGTTAAATGAATGAAGAGTTCAAGTGAAGTTTACCAAC 13418
 QY 1102 ACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161

Db 13419 AGCAATTGAGATGAGAGAGCTTAATCACAGGGTGAAGTGTAAAGAAAGCTACTG 13478
QY 1162 MAGAGATGACAAACACACTGTTTACAGCCAAAAATGAGTGTATATGCTAT-GGCTA 1220
Db 13479 ATGAGAGGAGACAGACATTAATACAGAACCACTGACAGAACTCAAAACACTCTGGAGA 13538
QY 1221 GCCTGAGNAGAGATCATCAACAGACATAAAGTGAAGGCTCTTATTTAGCCNTGG 1280
Db 13539 ACCTGGGTGAGAAATTTGCCACGACACAACTAGAGGGGCTCTGTTGGCCCTTG 13598
QY 1281 GTCAGTTCCAAATGAGTGTATGAGTGTCTGAT--GCTGACACACACGAGGGCTTGC 1338
Db 13599 GTCAGTTCCAGCATGCTTAAGAGAACTAATGAGTGGCTGATCATATCCGAAGAGTTGT 13658
QY 1339 TAAATGAGCAGAAACCTGTTGAGAGAACCTTAAAGCCATTGAATTGAATTGCCAAGC 1338
Db 13659 TAGATGCTCAGAGACCAATAGTGAAGAGCCAAAGATCATTTGAAGTGAAGCTCGCAAAGC 13718
QY 1399 ATCATGTGCTCCAAATGAGTGTATTTAGCCCATCAGTCCACAGTGAAGCCGTTAATTAAG 1458
Db 13719 ACCATGTCTTAAATAATGATTTTGGCTCATCAAGCCACAGTGAAGACATCAACAAAG 13778
QY 1459 CAGGAAATGATCTAATTTGAATCAAGTGCAGAGAGAAAGCAAGCACTTCAGAACAAAGC 1518
Db 13779 CTGGCAATGAGCTTCTTGAATCCAGTGTGAGATGATGCCAGAGCTTAAGAGCCGTT 13838
QY 1519 TAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAAAACAGAACAAAGAGCAGC 1578
Db 13839 TGGAAAGCCATGAAACCAATGCTGGAGAGTCAAGTTTACAGAAAAACAGAGAGAGGAGCAGC 13898
QY 1579 AGCTGGAATGATGCTGCGCCAGGCAAGGGTTCCATGGGGAATTGAGGATTTGCAGC 1638
Db 13899 AGCTTCACTCACTCTGACAGAGGCCAGGCTTCCACAGTGAATTTGAAGATTTCTCT 13958
QY 1639 AGTGGTGAAGTACACAGAGCGTATCTGTTGCAATTAACCGCTGGAGGTTTACCGG 1698
Db 13959 TGGAACTTACTAGAAATGAGAGCCAGCTTCTGCACTTAAGCCACAGAGAGACTTCTG 14018
QY 1699 AAACAGCCAGAGAGAGCTTAATGTCATATGAAAGTCTGTGCTGCTTTGAAGCTAAAG 1758
Db 14019 AAATGCTAGAGGAAACAGCTTATACACATATGAACTCTATTTCCAGCTGAAGCCAAAG 14078
QY 1759 AAGAAACATATAAGTCTGATGCAAGAAAGCCAGAGATGCTGCAAGATGCCAAAT 1818
Db 14079 AAGAGACTTATTAATCACTACTGCAAGGGCAAGCTCATGCTTTAAGCCGTGAGAGCT 14138
QY 1819 CTG 1821
Db 14139 CTG 14141

Search completed: January 12, 2005, 11:10:12
Job time : 8369 secs

THIS PAGE BLANK (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2005, 06:55:34 : Search time 5752 Seconds
(without alignments)
12866.672 Million cell updates/sec

Title: US-10-089-887-46

Perfect score: 2031
Sequence: 1 ttgcacatgcgcagcagaaca.....tngagagaccccaacgcgc 2031

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	633.6	31.2	1053	4	BG289408 602381438
2	623.6	30.7	676	5	BK472284 DKF2p686E
3	618.4	30.4	815	7	CK000929 AGENCOURT
4	561.6	27.7	764	7	CR549848 DKF2p469K
5	526.8	25.9	752	7	CK781134 UI-M-GV0-
6	486.4	23.9	800	5	BU701246 UI-M-EX0-
7	481	23.7	643	5	BK470591 DKF2p686M
8	478.2	23.5	589	7	CO700994 DG32-191b
9	469	23.1	826	6	CB164054 K-EST0225
10	468	23.0	866	6	BU139957 603134131
11	455.8	22.4	724	6	CB525847 UI-M-FY0-
12	449.8	22.1	799	6	CB722030 UI-M-GH0-
13	441.6	21.7	836	5	BU217687 603759087
14	440	21.7	793	5	BU368661 603597561
15	437.6	21.5	849	5	BQ215351 AGENCOURT
16	432.2	20.3	622	6	CA751356 UI-M-FO0-
17	413	20.3	511	1	AI889212 t250d03.x
18	406.2	20.0	749	7	CP743131 UI-M-GV0-
19	404.8	19.9	641	6	CB520499 UI-M-GI0-
20	403.2	19.9	748	6	CB520969 UI-M-GI0-
21	402.4	19.8	737	6	CB520605 UI-M-GI0-
22	400.8	19.7	703	4	BI731712 603353775
23	396.4	19.5	571	1	AI526522 uj39e02.Y
24	395.2	19.5	648	4	BM388627 UI-R-D20-

c	25	393	19.4	512	2	BE080914
	26	389	19.2	723	5	BU305460
	27	386	19.0	778	5	BU748872
	28	367.2	18.1	665	6	CD804239
	29	366.6	18.1	665	6	CD804438
	30	366.6	18.1	665	6	CD804591
	31	361	17.8	740	6	CB521821
c	32	358	17.6	361	4	BG949751
	33	353.2	17.4	486	4	BI345034
	34	345.6	17.0	490	6	CB548311
	35	340.2	16.8	719	5	BU477650
	36	335	16.5	805	6	CF284483
	37	333.8	16.4	446	2	BF871902
	38	328	16.1	331	2	BE003231
	39	322.8	15.9	533	2	AM475236
	40	321	15.8	427	6	CB794802
c	41	320.6	15.8	474	2	BF081412
	42	319.2	15.7	474	5	BK475908
	43	313.6	15.4	693	7	CN420731
	44	312.6	15.4	656	7	CR536640
	45	309.8	15.3	425	7	CK393268

ALIGNMENTS

RESULT 1
BG289408 1053 bp mRNA linear EST 21-FEB-2001
LOCUS 602381438F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4499097 5',
DEFINITION mRNA sequence.
BG289408
VERSION BG289408.1 GI:13045223
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strauberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10362 row: 1 column: 10
High quality sequence stop: 730.

FEATURES

source
1..1053
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4499097"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_93"
/note="Organ: Bladder; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

ORIGIN

Query Match 31.2%; Score 633.6; DB 4; Length 1053;
Best Local Similarity 95.7%; Pred. No. 9.8e-143;
Matches 748; Conservative 0; Mismatches 24; Indels 10; Gaps 9;

QY 181 GGGTTTTCATCCAGAGAAAGTATGTGGACCGGACCCCTTTACAGTCACTTAAGAA 240
DB 1 GGGTTTTCATCCAGAGAAAGTATGTGGACCGGACCCCTTTACAGTCACTTAAGAA 60
QY 241 GAGTCAAAAAGGTCG-TGTGGCAGTGAAGAACCTT-TCTCAATCA-CTCAGTTC 237
DB 61 GAGTCAAAAAGGTCGTCGTGACATGAGTGAAGCCATTATCTCAATCAAGTCACTTC 120
QY 298 CATGACAAAGATAGATCAGA-TCTTGGAGACC-TGGAAACGATCGTGAAGCTGAGGC 355
DB 121 CATGACAAAGATAGATCAGATTCTTGAAGCCGTCGAAACGATCTGGAAGCTGAGGC 180
QY 356 AGCCACCTCTATCTCTG-CAGAGTTGAGAAATCAAGAACAGATCACTGAAATTAAG 414
DB 181 AGCCACCTCTATCTCTGACAGAGTTGAGAAATCAAGAACAGATCACTGAAATTAAG 240
QY 415 AATGTGCACTAGACATGAAAAGCTACAGCCCTGTGATGAAACTCTTAACAGAGGGA 474
DB 241 AATGTGCACTAGACATGAAAAGCTACAGCCCTGTGATGAAACTCTTAACAGAGGGA 300
QY 475 GAGAAATGATGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCAAAAGCTGTTAG 534
DB 301 GAGAAATGATGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCAAAAGCTGTTAG 360
QY 535 GATAGCTTGACCAATGATGTTTCATTGAGAGAACATACACACTGTGTGAAGAGAG 594
DB 361 GATAGCTTGACCAATGATGTTTCATTGAGAGAACATACACACTGTGTGAAGAGAG 420
QY 595 GAGGCCAAATCTAGATGATGAGAGCTGACAAAAGTTCTGTGTGATCACTGTCA 654
DB 421 GAGGCCAAATCTAGATGATGAGAGCTGACAAAAGTTCTGTGTGATCACTGTCA 480
QY 655 TTGATAGTTACATTAAGATCTCAAGATTTTCACGGGACCTGGAAGATCTGGAAT 724
DB 481 TTGATAGTTACATTAAGATCTCAAGATTTTCACGGGACCTGGAAGATCTGGAAT 540
QY 715 GATCTTCACTAGTAAACCAACAGCAAGAACAGACACATTAAGGGAAGAAATAGAT 774
DB 541 GATCTTCACTAGTAAACCAACAGCAAGAACAGACACATTAAGGGAAGAAATAGAT 600
QY 775 GATCTTCACTAGTAAACCAACAGCAAGAACAGACACATTAAGGGAAGAAATAGAT 834
DB 601 GATCTTCACTAGTAAACCAACAGCAAGAACAGACACATTAAGGGAAGAAATAGAT 660
QY 835 GGGGAGCTGAT-AAACCATTTGTCAAGAGATATCATGATTAATTCAGCATGGG 893
DB 661 GGGGAGCTGAT-AAACCATTTGTCAAGAGATATCATGATTAATTCAGCATGGG 718
QY 894 ATTCTTAATTAAGCTTGAAGAACCG-GATTGACCACTTGAGAGCAATGACGTCGC 952
DB 719 GATCTTCACTAGTAAAGCTTGGGAGAGACCGCAATTGGCCAAATGGAAGAACGAGAGC 778
QY 953 GT 954
DB 779 GT 780

RESULT 2
BX472284 676 bp mRNA linear EST 04-SEP-2003
LOCUS BX472284
DEFINITION DKFZ686E12132_r1 686 (synonym: h1cc3) Homo sapiens cDNA clone
ACCESSION BX472284
VERSION BX472284.1 GI:3166552
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 676)
Bahr,A., Lauber,J., Mewes,H.W., Weill,B., Amlé,C., Osanger,A.,
Fodor,G., Han,W. and Wiemann,S.
EST (Bahr,A., Lauber,J., Mewes,H.W., Weill,B., et al.)

JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sequence available.
This clone (DKFZ686E12132) is available at the RZPD in Berlin.
Please contact the RZPD: Reesourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..676
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZ686E12132"
/dev_stage="adult"
/lab_host="DH10B"
/clone_11b="686 (synonym: h1cc3)"
/note="vector: pTriblex2; Site_1: SfiI; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match 30.7%; Score 623.6; DB 5; Length 676;
Best Local Similarity 97.6%; Pred. No. 2,4e-140;
Matches 652; Conservative 0; Mismatches 13; Indels 3; Gaps 2;
QY 1139 GCTTTTGTCAAAAGAAAGTAAACAGNAGAGTACAAACACTGTTCTAGACCAAAAT 1198
DB 9 GCTTTTGTCAAAAGAAAGTAAACAGNAGAGTACAAACACTGTTCTAGACCAAAAT 68
QY 1199 GGAATGATATTTGATAT-GCTTACCTGAGNAGATCATCAACAGACATTAAT 1257
DB 69 GGAATGAAATTTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 128
QY 1258 GGAAGGTGCTTAATTAAGCCNTGGGTGATGATGATGATGATGATGATGATGAT 1315
DB 129 GGAAGGTGCTTAATTAAGCCNTGGGTGATGATGATGATGATGATGATGATGAT 188
QY 1316 GCTGACACACACCGAGGGCTTCTAAGTACAGAAACCTGTTGAGAGACCTTAAGC 1375
DB 189 GCTGACACACACCGAGGGCTTCTAAGTACAGAAACCTGTTGAGAGACCTTAAGC 248
QY 1376 CATGAAATTTGAATCTTCCAGATCATGTGCTCCAAATGATGATTTAGCCCATGATC 1435
DB 249 CATGAAATTTGAATCTTCCAGATCATGTGCTCCAAATGATGATTTAGCCCATGATC 308
QY 1436 CAAGTGAAGACCGTTAATTAAGAGGAAATGATCTAATTAATGAATCAAGTGAAGAGAA 1495
DB 309 CAAGTGAAGACCGTTAATTAAGAGGAAATGATCTAATTAATGAATCAAGTGAAGAGAA 368
QY 1496 AGCAAGCACTTTCAGAACAGCTAGAGTTTAAATCAACCTGCGCAAAATGTTTTGA 1555
DB 369 AGCAAGCACTTTCAGAACAGCTAGAGTTTAAATCAACCTGCGCAAAATGTTTTGA 428
QY 1556 AAAAACAAGAAAGAGAGACAGCTGATGATGATGATGATGATGATGATGATGAT 1615
DB 429 AAAAACAAGAAAGAGAGACAGCTGATGATGATGATGATGATGATGATGATGAT 488
QY 1616 TGGCGAAATTAAGATTTGACAGATGATGATGATGATGATGATGATGATGATGAT 1675
DB 489 TGGCGAAATTAAGATTTGACAGATGATGATGATGATGATGATGATGATGATGAT 548
QY 1676 TAAACGCTGAGAGTTTACCGGAAACAGCAGAGAGAGCTTAAATGTCATATGGAAT 1735
DB 549 TAAACGCTGAGAGTTTACCGGAAACAGCAGAGAGAGCTTAAATGTCATATGGAAT 608
QY 1736 CTGTGCTGCTTTGAAGCTAAAGAGAAACATATTAAGTCTGATGACAGAAAGCCAGCA 1795
DB 609 CTGTGCTGCTTTGAAGCTAAAGAGAAACATATTAAGTCTGATGACAGAAAGCCAGCA 668

RESULT 3
CK000929
LOCUS CK000929 815 bp mRNA linear EST 26-NOV-2003
DEFINITION AGENCOURT_16363378 NIH_MGC_221 Homo sapiens cDNA clone IMAGE:30708013 5', mRNA sequence.
ACCESSION CK000929
VERSION CK000929.1 GI:38526963
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgsdb-remail.nih.gov
 Tissue Procurement: James Martin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: http://image.llnl.gov
 Plate: NDAMI073 row: 1 column: 14
 High quality sequence stop: 613.
FEATURES
 source
 1. 815
 location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30708013"
 /lab_host="DH10B Toba"
 /clone_lib="NIH_MGC_221"
 /note="Organ: mixed; Vector: pYX-Anc; Site_1: EcoRI; Site_2: NotI; Library is oligo-dT primed and directionally cloned. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to RNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Anc vector. Average insert size 4-5kb. Adaptors 5' (AATTCGCGACGAGG) 3' and 5'd (CGTCGCGCG) 3'. 3' linker sequence - GCGCGCGCTGAGACC T18. Sequencing primers 3' end: T3 promoter primer 5'd (ATTACCTCTACTTAAGGAG) 3'. 5' end: T7 promoter primer 5'd (TAATGACACTCACTATAGG) 3'. Library was constructed in the laboratory of M. Bento Soares. Note: this is a NIH_MGC library"
ORIGIN
 Query Match 30.4%; Score 618.4; DB 7; Length 815;
 Best Local Similarity 96.2%; Pred. No. 4,6e-139;
 Matches 685; Conservative 0; Mismatches 20; Indels 7; Gaps 5;
 1122 TGACATCATCAGCAGACGCTTTGGCTAAAGAAAGTAACAGNAGAGATGACAAACACACT 1181
 28 TGAATCATCAACAGCA-CTTTTGTAAAGAAAGTAACAGAGAGATGACAAACACACT 86
 1182 GTTCTAGACCCCAAAAATGAGNCTGATATTGATAT-GGCTAGCGCTGAGNAGAGAAATCATC 1240
 87 GTTCAAGACCCATTAATGGAATGAAATTTGATATGAGATAGCTGAGAGAGAAATCATC 146

1241 AACAGACGATAAATGAGAGGAGTCTATTATACCCNTGGGTCAGTTCCAAATGACCCCTG 1300
 147 AACAGACGATAAATGAGAGGAGTCTATTATAGCCTTGGTCAAGTTCCAAATGACCCCTG 206
 1301 TATGAGCTCCTG--CATGCTGACACACACCGAGGGCTTGTCTAAGTGAACGAAACCTGTT 1358
 207 GATGAGCTCCTGGCATGGCTGACACACCGAGGGCTTGTCTAAGTGAACGAAACCTGTT 266
 1359 GGAGAGACCCCTAAAGCATTTGAATTTGAATCTTGGCCAAAGCATCTATGCTCCAAATGAT 1418
 267 GGAGAGACCCCTAAAGCATTTGAATTTGAATCTTGGCCAAAGCATCTATGCTCCAAATGAT 326
 1419 GTATTAGCCCATCATGCTCCACAGTGAAGCCGTTATATTAAGACGAAATGATCTATTGAA 1478
 327 GTATTAGCCCATCATGCTCCACAGTGAAGCCGTTATATTAAGACGAAATGATCTATTGAA 386
 1479 TCAAGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1538
 387 TCAAGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 446
 1539 TGGCAAAATGTTTGGAAAAAACAAGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1598
 447 TGGCAAAATGTTTGGAAAAAACAAGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 506
 1599 CAGGCAAAAGGCTTCATAGGCAAAATGAGAGATTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1658
 507 CAGGCAAAAGGCTTCATAGGCAAAATGAGAGATTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 566
 1659 CGTCATCTGTTGGCATTAACCGCTGGAAGGTTTACCGGAAACAGGCAAGAGAGAGAGTT 1718
 567 CGTCATCTGTTGGCATTAACCGCTGGAAGGTTTACCGGAAACAGGCAAGAGAGAGAGTT 626
 1719 AATGTCATATGAGAGAGTGTGCTGCTCTTTGAAGACTTAAGAGAACTATTAAGTCTG 1778
 627 AATGTCATATGAGAGAGTGTGCTGCTCTTTGAAGACTTAAGAGAACTATTAAGTCTG 686
 1779 ATGCAAGAAAGGCGAGAGATGCTTGAAGAGATGCCCAAAATGTCAGAGAGAA 1830
 687 ATGCAAGAAAGGCGAGAGATGCTTGA--GATGCCCAAAATGTCAGAGAGAA 735
RESULT 4
CR549848
LOCUS CR549848 764 bp mRNA linear EST 12-JUL-2004
DEFINITION DKFZp69K061_r1 469 (synonym: pk1d1) Pongo pygmaeus cDNA clone DKFZp69K061 5', mRNA sequence.
ACCESSION CR549848
VERSION CR549848.1 GI:50243472
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
AUTHORS 1 (bases 1 to 764)
 Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R., Mewes, H.W., Well, B., Amid, C., Osengr, A., Fodor, G., Han, M., and Wiemann, S.
TITLE Pongo pygmaeus mRNA (Poustka, A., Albert, R., Moosmayer, P., et al.)
JOURNAL Unpublished (2004)
COMMENT Contact: MIPS
 MIPS
 Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp69K061) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gcf.de/projects/cdna/.
 Location/Qualifiers

FEATURES

```

source
1. .764
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp469k061"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469 (synonym: pkid1)"
/notes="Vector: pSport1_Sfi; Site_1: Sfi1A; Site_2: Sfi1B"

ORIGIN

Query Match 27.7%; Score 561.6; DB 7; Length 764;
Best Local Similarity 96.4%; Pred. No. 2.9e-125;
Matches 585; Conservative 0; Mismatches 20; Indels 2; Gaps 1;

QY 1283 CAGTTCACAAATGCTCTGTATGAGCTCTG--CATGTGACACACACCGAGGGCTTGCTA 1340
DB 1 CAGTTCACAAATGCTCTGTATGAGCTCTGACATGCTGACACACACCGAGGGCTTGCTA 60
QY 1341 AGTAGAGAGAAACCTGTTGAGAGAGACCTTAAGCCATTGAATTGAACCTTGCAAGCAT 1400
DB 61 AGTAGAGAGAAACCTGTTGAGAGAGACCTTAAGCCATTGAATTGAACCTTGCAAGCAT 120
QY 1401 CATGTGCTCCAAAATGATGATATTAGCCCATCAGTCCACAGTGGAAAGCCGTTAATAAGCA 1460
DB 121 CATGTGCTCCAAAATGATGATATTAGCCCATCAGTCCACAGTGGAAAGCCGTTAATAAGCA 180
QY 1461 GGAATATATCTAATTTGAATCAAGTGCAGGAGAGAAAGCAACCTTCGAAACAGCTA 1520
DB 181 GGAATATATCTAATTTGAATCAAGTGCAGGAGAGAAAGCAACCTTCGAAACAGCTA 240
QY 1521 GAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAAAACAGAAACAAAGAGAGAGAG 1580
DB 241 GAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAAAACAGAAACAAAGAGAGAGAG 300
QY 1581 CTGATGATGCTCTTGGCCAGCCAAAGGGTCTCATGCGCAAAATGAGATTGCGACAG 1640
DB 301 CTGATGATGCTCTTGGCCAGCCAAAGGGTCTCATGCGCAAAATGAGATTGCGACAG 360
QY 1641 TGGCTGACTGACACGAGAGGCTCATCTTGTGGCATCTAAACCGCTGGAGGTTTACCAGAA 1700
DB 361 TGGCTGACTGACACGAGAGGCTCATCTTGTGGCATCTAAACCGCTGGAGGTTTACCAGAA 420
QY 1701 ACAGCCAAAGAGAGAGCTTAATGTCATATGAAAGTCTGCTGCTTTGAAGCTTAAGAA 1760
DB 421 ACAGCCAAAGAGAGAGCTTAATGTCATATGAAAGTCTGCTGCTTTGAAGCTTAAGAA 480
QY 1761 GAAACATATTAAGAGTCTGATGCAAGAAAGCCAGACAGATGCTTGCAGAGATGCCAAATCT 1820
DB 481 GAAACATATTAAGAGTCTGATGCAAGAAAGCCAGACAGATGCTTGCAGAGATGCCAAATCT 540
QY 1821 GCAGAGCAAAATTTTGCCCAAGACATTAATTAACCTGAAAAAATTTTTTTTTTTTTTTT 1880
DB 541 GCAGAGCAAAATTTTGCCCAAGACATTAATTAACCTGAAAAAATTTTTTTTTTTTTTTT 600
QY 1881 AAAANAC 1887
DB 601 ACCAACC 607

RESULT 5
CK781134 752 bp mRNA linear EST 23-FEB-2004
LOCUS CK781134
DEFINITION U1-M-GVO-c1n-d-16-0-U1.r1 NIH_BMAP_GVO Mus musculus cDNA clone
IMAGE:30616311 5', mRNA sequence.
ACCESSION CK781134
VERSION CK781134.1 GI:42746812
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE
1 (bases 1 to 752)
AUTHORS NIH-WGC http://imgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution Information can be found at
http://genome.uiowa.edu/distribution/mouseefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.

FEATURES
source
Location/Qualifiers
1..752
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAG:30616311"
/tissue_type="whole brain"
/dev_stage="1,5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GVO"
/notes="Organ: Brain; Vector: pYX-Abs; Site_1: Ecor I;
Site_2: Not I; The library was constructed according
to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
Ecor I adaptor, digested with NotI and then cloned
directionally into pYX-Abs vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAATCGAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match 25.9%; Score 526.8; DB 7; Length 752;
Best Local Similarity 82.2%; Pred. No. 8.2e-117;
Matches 617; Conservative 0; Mismatches 130; Indels 4; Gaps 1;

QY 99 TGGTGAGTTGATAGCTGACACAAAGCTCATATATGATTAAGTGAACAAACTGGCCA-- 156
DB 1 TTGGAAGACTGATAGCTGACACAAAGCTCATATATGATTAAGTGAACAAAGCTGGCCAC 60
QY 157 --CAGTACTGATGAGCCCTGGGAAAGGCTTTCTTATCCAAAGAAAGTATGCGACGGC 214
DB 61 AGCTATTGGAACTGAGCCCAAGGAGGATTTATTCAGAAAGATATGCGACGGC 120
QY 215 ACACCCCTTACAGTCACTTAAGAAAGATGTCAAAAAGGCTGCGACCTGATGAAG 274
DB 121 ACACCCCTTACAGTCACTTAAGAAAGATGTCAAAAAGGCTGCGACCTGATGAAG 180
QY 275 CCATTTCTCAATCACTCACTGATGACATGACAAATGATGATGCTTGAAGACCTGGAAC 334
DB 181 CCATTTCTCAATCACTCACTGATGACATGACAAATGATGATGCTTGAAGACCTGGAAC 240
QY 335 GCATCTGGAAGAGTCTGAGGCGAGCCACCTCTATCTCTGCAAGGTTGGAAGATTAAG 394
DB 241 GCATCTGGAAGAGTCTGAGGCGAGCCACCTCTATCTCTGCAAGGTTGGAAGATTAAG 300
QY 395 AACAGATCAGTGAATAATGAATGTCAGTGAACATGGAAGAAAGCTACAGCCGTTATG 454
DB 301 AACAGATGTCGAATAATGAATGTCAGTGAACATGGAAGAAAGCTACAGCCGTTATG 360
QY 455 AAACCTTTAAACAGAGGAGAGAAATGATGCTGATGATCTGGGGGACATGATTAAGACA 514

```


Db 361 AGACTCTGAGCAGAGGGAGAGAGATGATCGCAGATCCGAGGACCTGAAAAAGAC 420
Qy 515 TATCTGCCAAAGCTCTTCAAGATAAGCTTGACCAATGTTTCTTTGGAGAACTAC 574
Db 421 TATCTGCCAGAGCTGTACAGATTAAGTTGACCAATGTTTCTTTGGAGAACTAC 480
Qy 575 ACACACTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 634
Db 481 ACACCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 635 TCTGTGTGATCAGATGTCATGATGATGATGATGATGATGATGATGATGATGAT 694
Db 541 TCTGTGTGATCAGATGTCATGATGATGATGATGATGATGATGATGATGATGAT 600
Qy 695 ACCTGGAAGATCTCTGGAATTTGATCTTCACTAAGTAAACACAGAGAGAGAGAG 754
Db 601 ACCTGGAAGATCTCTGGAATTTGATCTTCACTAAGTAAACACAGAGAGAGAG 660
Qy 755 CCATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 814
Db 661 CAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy 815 CTGAATCTATGCGGCGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 845
Db 721 CAGAGCTCATCGCTGCGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751

RESULT 6
LOCUS BU701246 800 bp mRNA linear EST 15-JUL-2003
DEFINITION UI-M-EX0-bwy-a-04-0-UI.r1 NIH_BMAP_EX0 Mus musculus cDNA clone
ACCESSION BU701246 IMAGE:5704587 5', mRNA sequence.
VERSION BU701246.1 GI:23624770
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIN at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: PYX-5.
Location/Qualifiers
1. 800
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5704587"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_id="NIH BMAP EX0"
/note="Organ: brain; Vector: pYX-asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction.

ligated with Ecor I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA
tail, is GTCGTGGA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 23.9%; Score 486.4; DB 5; Length 800;
Best Local Similarity 81.3%; Pred. No. 5.4e-107;
Matches 610; Conservative 0; Mismatches 135; Indels 5; Gaps 4;

Qy 1095 TATCAACAGCAGATAG-AGTGTGACAGTCACTATCAACGAGAGCGCTTTGCTAAAGAA 1153
Db 2 TATCAACAAACAGATAGAAATGAAAGGCTGAAACATCAAGCAGA-AGTGTGCTGAAGAA 60
Qy 1154 AGTAAAGNAGAGAGTGAACAAACACGCTTCTAGACCAAAATGAGATGATATGAT 1213
Db 61 AGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 120
Qy 1214 AT-GGCTAGCCTGAGAGAGAGATCATCAACAGAGATTAATGAGAGGCTCTAAT 1272
Db 121 CTGGAGAGAGAGAGAGAGAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 1273 AGCCTGAGTCACTTCAACATGCTGTATAGCTCT--GCATGCTGACACACCGGA 1330
Db 181 AGCAGTGGGTGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 1331 GGGCTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1330
Db 241 AGGCTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Qy 1391 TGCCAGAGATGATGCTCTCAAAATGATGATGATGATGATGATGATGATGATGATG 1450
Db 301 GGTAAAGATGATGCTCTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 360
Qy 1451 TAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1510
Db 361 CAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 1511 GAACAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1510
Db 421 GTAAACCTCAGATTTTCAATCAACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 1571 GAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1630
Db 481 AAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 1631 TTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1690
Db 541 TTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 1691 TTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1750
Db 601 TCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 1751 AGCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1810
Db 661 CATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy 1811 CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1840
Db 721 TCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 750

RESULT 7
LOCUS BX470991 643 bp mRNA linear EST 04-SEP-2003
DEFINITION DK22686M17123 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION BX470991 DK22686M17123 5', mRNA sequence.

VERSION	BK470991.1	GI:31665308
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 643)	
AUTHORS	Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Aml, C., Oeanger, A., Fobo, G., Han, M., and Wiemann, S.	
TITLE	EST (Bahr, A., Lauber, J., Mewes, H.W., Weil, B., et al.)	
JOURNAL	Unpublished (2003)	
COMMENT	Contact: MIPS	

FEATURES
Ingolstaedter Landstr.1,D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert
clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Olagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No **at** sequence available.
This clone (DKFZp686k17123) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

FEATURES	Location/Qualifiers
source	1. .643

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZ686M17123"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="686 (synonym: hicc3)"
/name="Vector: pRipIDex2; site_1: SfilA; site_2: SfilB;
CDNA-collection"

```

ORIGIN

Query Match	23.7%	Score 481	DB 5	Length 643
Best Local Similarity	91.0%	Pred. NO. 1.1e-105		
Matches 588	Conservative 0	Mismatches 45	Indels 13	Gaps 7

QY	722	CAGTAGTAAACAACAGCAAGACAGAGACCATAGAGGAAGAAATAGTGGACATC	781
Db	1	CAGTAGTAAACAACAGCAAGACAGAGACCATAGAGGAAGAAATAGTGGACATC	60
QY	782	AGAGAGAGCTGGATATAGTATTTAACTAGGTTCTGAACTCATGGCGCATGTGGAGGC	841
Db	61	AGAGAGAGCTGGATATAGTATTTAACTAGGTTCTGAACTCATGGCGCATGTGGAGGC	120
QY	842	CTGATTAACCCATTTGTCAAGAAAGATATACATAGTTAAATTCAAGCGATGGGATTCCTTA	901
Db	121	CTGATTAACCCATTTGTCAAGAAAGATATAGTTAAATTCAAGC-ATGGATATCTTA	179
QY	902	AATTAAGCTTGGAAAGACCGGATTTGACCAACTTGAG-----AGCAATGCAGCTGCCGTC	955
Db	180	AATTAAGCTTGGAAAGACCGGATTTGACCAACTTGAGAGGCAATGCAGCTGCCGTTAG	239
QY	956	AGTACAGATGGACTGCAGAGCGGTATTTTTCGGGTATGATTCAGAGGTGTAAAGTGGCT	1015
Db	240	TACCAAGATGGACTGCAGAGCGGTATTTTTCGGGTATGATTCAGAGGTGTAAATTTACT	299
QY	1016	TCAATGTCCTCAATTTGGAACAGCATCTTCGAATCTGCAAGACGAGATTTGATGTCTGA	1075
Db	300	TCAATGTCCTCAATTTGGAACAG-ATCTGAAACTGTCAAGCAGCAGATTTGAAGAGCTTAA	358
QY	1076	GCAATTTTAAGTCTGAGGCTTATCAACAGCAAGTAG-AGTGTGACGACTGACTCATCAGC	1134
Db	359	GCAATTTTAAGTCTGAGGCTTATCAACAGCAAGTGAAGAAAGATGAAATCATCAAGC	418
QY	1135	AGAGCGCTTTTGTCTAAAGAAATTAACAGAAAGAGATGACAAACACACTGTTCTAAGCCAA	1194
Db	419	AGA-GCTTTTGTCTAAAGAAATTAACAGAAAGAGATGACAAACACACTGTTCTAAGCCAT	477
QY	1195	AAATAGGATCTGATATTGATAT-GGCTAGCCTTGAGAGNAGAGATCATCAACAGACAGATA	1253

Db	478	TTATGGAACCTGAATTGATATGAGATAGCCTGGAGGAGAGAAATCATCAACAGACAGCATTA	537
Qy	1254	AACCTGGAGGGGCGCTCTATTTAGCCNVTGGGTACGTTCCAAACATGGCCCTGATATAGCTCCNG-	1312
Db	538	AACCTGGAGGGGCGCTCTATTTAGCCTTGGGTCAGTTCCAAACATGGCCCTGATATAGCTCTGG	597
Qy	1313	-CATGCTGACACACACCGAGGGGCTTGCTAATGTAGCAAAACCTGT	1357
Db	598	CATGCTGACACACACCGAGGGGCTTGCTAATGTAGCAAAACCTGT	643

RESULT 8				
CO700994/c				
LOCUS	CO700994	589 bp	mRNA	linear
DEFINITION	DG32-191b14	DG32-1	Canis familiaris	cDNA 3', mRNA sequence.

**DOORCE
ORGANISM**

Canis familiaris (dog)

REFERENCE

1 (bases 1 to 589)

1

Henrich, J. and Loebbert, R.

JOURNAL

Unpublished (2004)

LION bioscience AG

Tel: +49 6221 4038 150

Email: Thomas.Schluter@lionbioscience.com

SOURCE:

1: .589

ORIGIN

Query Match	23.5%	Score	478.2	DB	7	Length	589
Best Local Similarity	89.7%	Pred. No.	5e-105				
Matches	524	Conservative	0	Mismatches	59	Indels	1
						Gaps	1

Qy	176	GGGAAGCCTTTCTATCCAAAGAAAGATGTGGAGCGACACCCCTTTACGTACAATTA	235
Db	587	GGTGAAGCGCTTTCTATCCAAAGAAAGTACGTGACGTACACCC-TTACAGTCAAAATTA	529
Qy	236	AAGAAGATGTCAAAAACCGTGTGTGGCACTGTATGTAAGCCATTTCTCAATCAACTAGT	295
Db	528	AAGAAGATGTCAAAAACCGAGCGGTGCATTGGATGAAGCATTTTCTCAATCTACTAGT	469
Qy	296	TCCATGCAAGATATGATCAGTCTCTTGAAGCCTGTGAACGCATGTGTGAACGTCTGAGGC	355
Db	468	TCCACGACAAAGATAGACCAATATCTTGAAGCCGTGAACGCATGTGTGAGCGTTTGAAGGC	409
Qy	356	AGCCACCCCTATCTCTGCAAGGTTGTGAAGATCAAGSAAACAAATCAGTGAATAATAAGA	415
Db	408	AGCCGCTTGGATCTCAGCTGAGGTGAAAAGATTTAAGAGCAAAATCATGTAGAAATTAAGA	349
Qy	416	ATGTGTCAGTAGACATGSAAAAGCTAGCGCGTGTATGAACTCTTAAACAGAGGGGAG	475
Db	348	ATGTGTCAGTATGATATGGAAGAGTAAACCATTTGTATGAAACCTTTAAAAAGGGGGAG	289
Qy	476	AGGAATGATTTGCTAGATCTGGGGGAGCTGATAAAGCATATCTGCCAAAGCTGTTACGG	535

Db 288 AGAATGATGCAAGATCTGAGGTAAGCAAGACATATCTGCAGAAAGCTGTTCAG 229

Qy 536 ATAGCTTGACCAATGGTTTTCATTTGGGAGAACATCACACTGTGTGAAAGAGGG 595

Db 228 ATAGCTTGACCAATGGTTTTCATTTGGGAGAACATCACACTGTGTGAAAGAGGG 169

Qy 596 AAGCAAACTCTGATGATGAGCTAGCAGAAAAGTTCTGGTGTGATCACATGTGAT 655

Db 168 AAGCAAACTCTGATGATGAGCTAGCAGAAAAGTTCTGGTGTGATCACATGTGAT 109

Qy 656 TGATAGTTCATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 715

Db 108 TGATAGTTCATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 49

Qy 716 ATCTTCAGTGTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 759

Db 48 ATCCCTCAGTGTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 5

RESULT 9
CB164054/c 469 bp mRNA linear EST 30-JAN-2003

LOCUS K-EST0225079 L17N670205n1 Homo sapiens cDNA clone

DEFINITION L17N670205n1-37-C03 5', mRNA sequence.

ACCESSION CB164054

VERSION CB164054.1 GI:28150180

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 469)
Kam,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

AUTHORS

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsun@gmail.kr
Plate: 37 row: C column: 03
High quality sequence stop: 469.

FEATURES
source
1. .469
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L17N670205n1-37-C03"
/sex="F"
/lab_host="Top10P"
/clone_lib="L17N670205n1"
/note="Organ: Liver; Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; The library was contributed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."

ORIGIN
Query Match 23.1%; Score 469; DB 6; Length 469;
Best Local Similarity 100.0%; Pred. No. 8.1e-103; Indels 0; Gaps 0;
Matches 469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1369 CTAAGCCATTGAATTAAGTTCGCAAGCATCTGCTCCAAATATGATTAGCCC 1428

Db 469 CTAAGCCATTGAATTAAGTTCGCAAGCATCTGCTCCAAATATGATTAGCCC 410

Qy 1429 ATGATCCAGAGTGAAGCGCTTAATTAAGCAGAAATGATCTAATTGAATCAAGTGAG 1488

Db 409 ATGATCCACAGTGAAGCCGTTAATTAAGCAGAAATGATCTAATTGAATCAAGTCAG 350

Qy 1489 GAGAAAGACAGAGACCTTTCAGAAACAAGCTTAAGAGTTTAAATCAACGCTGGCAAAATG 1548

Db 349 GAGAAAGACAGAGACCTTTCAGAAACAAGCTTAAGAGTTTAAATCAACGCTGGCAAAATG 290

Qy 1549 TTTTGAAAAAACAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1608

Db 289 TTTTGAAAAAACAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 230

Qy 1609 GGTTCATGAGGAAATTAAGAGATTTGAGAGAGTGTGACTGACACGAGAGGTCATGCT 1668

Db 229 GGTTCATGAGGAAATTAAGAGATTTGAGAGAGTGTGACTGACACGAGAGGTCATGCT 170

Qy 1669 TGGCATCTAAACCCCTGGAGGTTTACCGGAACAGCCAGAGACGATTATGTCATA 1728

Db 169 TGGCATCTAAACCCCTGGAGGTTTACCGGAACAGCCAGAGACGATTATGTCATA 110

Qy 1729 TGGAGTCTGCTGCTGCTTGAAGCTAAAGAAACATTAAGAGTGTGAGAGAG 1788

Db 109 TGGAGTCTGCTGCTGCTTGAAGCTAAAGAAACATTAAGAGTGTGAGAGAG 50

Qy 1789 GCCAGAGATGCTTGAAGATGCCCAAAATCTGCAGAGACAAATATTGA 1837

Db 49 GCCAGAGATGCTTGAAGATGCCCAAAATCTGCAGAGACAAATATTGA 1

RESULT 10
BU139957 826 bp mRNA linear EST 25-NOV-2002

LOCUS BU139957

DEFINITION 603134131.F1 CSEBCHL24 Gallus gallus cDNA clone CHEST115n8 5', mRNA sequence.

ACCESSION BU139957

VERSION BU139957.1 GI:25354439

KEYWORDS EST.

SOURCE Gallus gallus

ORGANISM Gallus gallus (chicken)

REFERENCE Boardman,P.E., Sans-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

AUTHORS

TITLE 1245392

JOURNAL Contact: Simon Hubbard

MEDLINE Department of Biomolecular Sciences

PUBMED University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 10D, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1. .826
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST115n8"
/dev_stage="16 day embryo"
/lab_host="DH10B"
/clone_lib="CSEBCHL24"
/note="Organ: brain; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)
[stratagene] vector to accommodate cDNA produced with the T-primed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BglI and

ORIGIN

BamHI sites [5'ggccgctgacgcccgcgagtcggaataaaag] [5'aattctttcttcgagatccggggctgcacgc]"

Query Match 23.0%; Score 468; DB 5; Length 826;
 Best Local Similarity 79.1%; Pred. No. 1.6e-102;
 Matches 653; Conservative 0; Mismatches 164; Indels 9; Gaps 8;

QY 996 TTGACAGTGGTAGTTCGCTTCAATGCTCCAAATGGAAACAGACATCGAACTGTCAG 1055
 DB 2 TTGCTGGACATCAAGCTAGATCTATGTCCCGCTTGAAC-GGATCTGAGACAGTGAAG 60
 QY 1056 CACGACATGATGTCTAGAGCAATTTAACTGAGGCTTATCAACAGCAGATAG-AGTG 1114
 DB 61 CAACAACTGAGGAGACTTAAAGCAATTTAAACAGAGCTTATCAGACAGCAGATAGAATG 120
 QY 1115 GTACGACTACTATACCGAGACGCTTTGCTAAAGAAAGTAAACAGNAGAGTACAA 1174
 DB 121 GAAAGCTGAACATCAAGCAGA-ACATATTGCTGAAGAGTAAACAAGGAAAGTACAA 179
 QY 1175 ACACACTGTTCTAGACCCCAAAAATGAGCTGATATTGAT-GGCTAGCTGGAGNAGAG 1233
 DB 180 GCACACTGTTCAAGACCCCTCTCAGAGCTCAGAAATATGTGGACAGCCTGGAGGAAAA 239
 QY 1234 AATCATCAACAGACAGATAACTGAGGGTGTCTATTAGCCNTGGCTCAGTTCCAA 1293
 DB 240 AATTATTAATAGACAGACAGCAAGCTGGAGGGTGTCTATTAGCTGGGAGAGTTCCAGCA 239
 QY 1294 TGGCTCTGTAGACTCT--GCATGCTGACACACCGAGGGCTTGT-AGGTGAGCAGA 1350
 DB 300 TGGCTGTGATGATTTACTTACGTGCTGACGCAACAGAGACTTGTGTAAGTGAACAGA 359
 QY 1351 AACCTGTGAGAGAGACCTTAAGCCATTGAAATTTGCAAGCATCATGTGCTC 1410
 DB 360 AACCTGGGTGAGAGACCCCAAGCTTTTAAATTTGCACTTGGCAACATCATGTGCTAC 419
 QY 1411 AAAATGATATTAGCCCATCACTGCAACAGTGAAGCCGTTAAATAAGCAGAAATGATC 1470
 DB 420 AGATGATGTTTATAGCTCATCAGCTTACAGTAGAGAGTTTAAAGAAAGAGTGAAC 479
 QY 1471 TAATTGAATCAATGACAGAGAAAGCAAGCAACCTTCAAGCAACAGCTAGAGTTTAA 1530
 DB 480 TGAATGATCAATGTGCTGTTGAAGAGCAAGTAAATTTACGAGCAAGTGTGAACCTTCTGA 539
 QY 1531 ATCAAGCTGGCAAAATGTTTGGAAAAAACAAGAACAAAGAGACAGCTGAGTGGT 1590
 DB 540 ATCAAGCTGGCAAAATGTTTGGAAAAAACAAGAACAAAGAGACAGCTGAGTGGT 599
 QY 1591 CCTTGGCCAGGCAAGAGGTTCCATGCGCAAAATTTGAGATTTGACAGTGGCTGACTG 1650
 DB 600 CCTTATCCAGGCGCAAGGTTTCCATGTGTGAAGTTAGAGTATGACAGCATGTGCTGACG 659
 QY 1651 ACAGGAGCGTCACTGTGTTGGCATTTAAACGCTGGAGGTTTACCGGAAAACAGCCAGG 1710
 DB 660 ACACTAAACGTCAACTGTTGGCATCAAAACCTGTTGAGGCTTACAGAAACAGCAAGG 719
 QY 1711 AGAGAGCTAA-TGTCCATATGAGAGTGTGCTGCC-TTGAAGCTTAAAGAAACATTA 1768
 DB 720 AACAGCTTAATTAACCATATGAGAACTTGTGCTGCTTTTGAAGCCAAAGAGACATA 779
 QY 1769 TAAAGTCTGATGACAGAAAGGACAGATGATTTGCAAGTGGCCA 1814
 DB 780 TAGGTGTCTTAATGACAGAACGCGCAGAGATGCTTGCAGAGTGCCTCA 825

RESULT 11
 C8525847 724 bp mRNA linear EST 09-JUN-2003
 LOCUS UI-M-FY0-cfe-p-07-0-UI-r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 DEFINITION IMAGE:6848600 5', mRNA sequence.
 ACCESSION C8525847
 VERSION C8525847.1 GI:29359318
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 724)
 REFERENCE NIH-MGC <http://mgc.ncl.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: c8525847-ncl.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 DNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at <http://genome.uiowa.edu/distribution/mouse1.html>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 Seq primer: PYX-5,
 Location/Qualifiers
 1..724
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6848600"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5, 14.5, 16.5, 17.5, 18.5, 19.5, 20.5, 21.5, 22.5, 23.5, 24.5, 25.5, 26.5, 27.5, 28.5, 29.5, 30.5, 31.5, 32.5, 33.5, 34.5, 35.5, 36.5, 37.5, 38.5, 39.5, 40.5, 41.5, 42.5, 43.5, 44.5, 45.5, 46.5, 47.5, 48.5, 49.5, 50.5, 51.5, 52.5, 53.5, 54.5, 55.5, 56.5, 57.5, 58.5, 59.5, 60.5, 61.5, 62.5, 63.5, 64.5, 65.5, 66.5, 67.5, 68.5, 69.5, 70.5, 71.5, 72.5, 73.5, 74.5, 75.5, 76.5, 77.5, 78.5, 79.5, 80.5, 81.5, 82.5, 83.5, 84.5, 85.5, 86.5, 87.5, 88.5, 89.5, 90.5, 91.5, 92.5, 93.5, 94.5, 95.5, 96.5, 97.5, 98.5, 99.5, 100.5, 101.5, 102.5, 103.5, 104.5, 105.5, 106.5, 107.5, 108.5, 109.5, 110.5, 111.5, 112.5, 113.5, 114.5, 115.5, 116.5, 117.5, 118.5, 119.5, 120.5, 121.5, 122.5, 123.5, 124.5, 125.5, 126.5, 127.5, 128.5, 129.5, 130.5, 131.5, 132.5, 133.5, 134.5, 135.5, 136.5, 137.5, 138.5, 139.5, 140.5, 141.5, 142.5, 143.5, 144.5, 145.5, 146.5, 147.5, 148.5, 149.5, 150.5, 151.5, 152.5, 153.5, 154.5, 155.5, 156.5, 157.5, 158.5, 159.5, 160.5, 161.5, 162.5, 163.5, 164.5, 165.5, 166.5, 167.5, 168.5, 169.5, 170.5, 171.5, 172.5, 173.5, 174.5, 175.5, 176.5, 177.5, 178.5, 179.5, 180.5, 181.5, 182.5, 183.5, 184.5, 185.5, 186.5, 187.5, 188.5, 189.5, 190.5, 191.5, 192.5, 193.5, 194.5, 195.5, 196.5, 197.5, 198.5, 199.5, 200.5, 201.5, 202.5, 203.5, 204.5, 205.5, 206.5, 207.5, 208.5, 209.5, 210.5, 211.5, 212.5, 213.5, 214.5, 215.5, 216.5, 217.5, 218.5, 219.5, 220.5, 221.5, 222.5, 223.5, 224.5, 225.5, 226.5, 227.5, 228.5, 229.5, 230.5, 231.5, 232.5, 233.5, 234.5, 235.5, 236.5, 237.5, 238.5, 239.5, 240.5, 241.5, 242.5, 243.5, 244.5, 245.5, 246.5, 247.5, 248.5, 249.5, 250.5, 251.5, 252.5, 253.5, 254.5, 255.5, 256.5, 257.5, 258.5, 259.5, 260.5, 261.5, 262.5, 263.5, 264.5, 265.5, 266.5, 267.5, 268.5, 269.5, 270.5, 271.5, 272.5, 273.5, 274.5, 275.5, 276.5, 277.5, 278.5, 279.5, 280.5, 281.5, 282.5, 283.5, 284.5, 285.5, 286.5, 287.5, 288.5, 289.5, 290.5, 291.5, 292.5, 293.5, 294.5, 295.5, 296.5, 297.5, 298.5, 299.5, 300.5, 301.5, 302.5, 303.5, 304.5, 305.5, 306.5, 307.5, 308.5, 309.5, 310.5, 311.5, 312.5, 313.5, 314.5, 315.5, 316.5, 317.5, 318.5, 319.5, 320.5, 321.5, 322.5, 323.5, 324.5, 325.5, 326.5, 327.5, 328.5, 329.5, 330.5, 331.5, 332.5, 333.5, 334.5, 335.5, 336.5, 337.5, 338.5, 339.5, 340.5, 341.5, 342.5, 343.5, 344.5, 345.5, 346.5, 347.5, 348.5, 349.5, 350.5, 351.5, 352.5, 353.5, 354.5, 355.5, 356.5, 357.5, 358.5, 359.5, 360.5, 361.5, 362.5, 363.5, 364.5, 365.5, 366.5, 367.5, 368.5, 369.5, 370.5, 371.5, 372.5, 373.5, 374.5, 375.5, 376.5, 377.5, 378.5, 379.5, 380.5, 381.5, 382.5, 383.5, 384.5, 385.5, 386.5, 387.5, 388.5, 389.5, 390.5, 391.5, 392.5, 393.5, 394.5, 395.5, 396.5, 397.5, 398.5, 399.5, 400.5, 401.5, 402.5, 403.5, 404.5, 405.5, 406.5, 407.5, 408.5, 409.5, 410.5, 411.5, 412.5, 413.5, 414.5, 415.5, 416.5, 417.5, 418.5, 419.5, 420.5, 421.5, 422.5, 423.5, 424.5, 425.5, 426.5, 427.5, 428.5, 429.5, 430.5, 431.5, 432.5, 433.5, 434.5, 435.5, 436.5, 437.5, 438.5, 439.5, 440.5, 441.5, 442.5, 443.5, 444.5, 445.5, 446.5, 447.5, 448.5, 449.5, 450.5, 451.5, 452.5, 453.5, 454.5, 455.5, 456.5, 457.5, 458.5, 459.5, 460.5, 461.5, 462.5, 463.5, 464.5, 465.5, 466.5, 467.5, 468.5, 469.5, 470.5, 471.5, 472.5, 473.5, 474.5, 475.5, 476.5, 477.5, 478.5, 479.5, 480.5, 481.5, 482.5, 483.5, 484.5, 485.5, 486.5, 487.5, 488.5, 489.5, 490.5, 491.5, 492.5, 493.5, 494.5, 495.5, 496.5, 497.5, 498.5, 499.5, 500.5, 501.5, 502.5, 503.5, 504.5, 505.5, 506.5, 507.5, 508.5, 509.5, 510.5, 511.5, 512.5, 513.5, 514.5, 515.5, 516.5, 517.5, 518.5, 519.5, 520.5, 521.5, 522.5, 523.5, 524.5, 525.5, 526.5, 527.5, 528.5, 529.5, 530.5, 531.5, 532.5, 533.5, 534.5, 535.5, 536.5, 537.5, 538.5, 539.5, 540.5, 541.5, 542.5, 543.5, 544.5, 545.5, 546.5, 547.5, 548.5, 549.5, 550.5, 551.5, 552.5, 553.5, 554.5, 555.5, 556.5, 557.5, 558.5, 559.5, 560.5, 561.5, 562.5, 563.5, 564.5, 565.5, 566.5, 567.5, 568.5, 569.5, 570.5, 571.5, 572.5, 573.5, 574.5, 575.5, 576.5, 577.5, 578.5, 579.5, 580.5, 581.5, 582.5, 583.5, 584.5, 585.5, 586.5, 587.5, 588.5, 589.5, 590.5, 591.5, 592.5, 593.5, 594.5, 595.5, 596.5, 597.5, 598.5, 599.5, 600.5, 601.5, 602.5, 603.5, 604.5, 605.5, 606.5, 607.5, 608.5, 609.5, 610.5, 611.5, 612.5, 613.5, 614.5, 615.5, 616.5, 617.5, 618.5, 619.5, 620.5, 621.5, 622.5, 623.5, 624.5, 625.5, 626.5, 627.5, 628.5, 629.5, 630.5, 631.5, 632.5, 633.5, 634.5, 635.5, 636.5, 637.5, 638.5, 639.5, 640.5, 641.5, 642.5, 643.5, 644.5, 645.5, 646.5, 647.5, 648.5, 649.5, 650.5, 651.5, 652.5, 653.5, 654.5, 655.5, 656.5, 657.5, 658.5, 659.5, 660.5, 661.5, 662.5, 663.5, 664.5, 665.5, 666.5, 667.5, 668.5, 669.5, 670.5, 671.5, 672.5, 673.5, 674.5, 675.5, 676.5, 677.5, 678.5, 679.5, 680.5, 681.5, 682.5, 683.5, 684.5, 685.5, 686.5, 687.5, 688.5, 689.5, 690.5, 691.5, 692.5, 693.5, 694.5, 695.5, 696.5, 697.5, 698.5, 699.5, 700.5, 701.5, 702.5, 703.5, 704.5, 705.5, 706.5, 707.5, 708.5, 709.5, 710.5, 711.5, 712.5, 713.5, 714.5, 715.5, 716.5, 717.5, 718.5, 719.5, 720.5, 721.5, 722.5, 723.5, 724.5, 725.5, 726.5, 727.5, 728.5, 729.5, 730.5, 731.5, 732.5, 733.5, 734.5, 735.5, 736.5, 737.5, 738.5, 739.5, 740.5, 741.5, 742.5, 743.5, 744.5, 745.5, 746.5, 747.5, 748.5, 749.5, 750.5, 751.5, 752.5, 753.5, 754.5, 755.5, 756.5, 757.5, 758.5, 759.5, 760.5, 761.5, 762.5, 763.5, 764.5, 765.5, 766.5, 767.5, 768.5, 769.5, 770.5, 771.5, 772.5, 773.5, 774.5, 775.5, 776.5, 777.5, 778.5, 779.5, 780.5, 781.5, 782.5, 783.5, 784.5, 785.5, 786.5, 787.5, 788.5, 789.5, 790.5, 791.5, 792.5, 793.5, 794.5, 795.5, 796.5, 797.5, 798.5, 799.5, 800.5, 801.5, 802.5, 803.5, 804.5, 805.5, 806.5, 807.5, 808.5, 809.5, 810.5, 811.5, 812.5, 813.5, 814.5, 815.5, 816.5, 817.5, 818.5, 819.5, 820.5, 821.5, 822.5, 823.5, 824.5, 825.5, 826.5, 827.5, 828.5, 829.5, 830.5, 831.5, 832.5, 833.5, 834.5, 835.5, 836.5, 837.5, 838.5, 839.5, 840.5, 841.5, 842.5, 843.5, 844.5, 845.5, 846.5, 847.5, 848.5, 849.5, 850.5, 851.5, 852.5, 853.5, 854.5, 855.5, 856.5, 857.5, 858.5, 859.5, 860.5, 861.5, 862.5, 863.5, 864.5, 865.5, 866.5, 867.5, 868.5, 869.5, 870.5, 871.5, 872.5, 873.5, 874.5, 875.5, 876.5, 877.5, 878.5, 879.5, 880.5, 881.5, 882.5, 883.5, 884.5, 885.5, 886.5, 887.5, 888.5, 889.5, 890.5, 891.5, 892.5, 893.5, 894.5, 895.5, 896.5, 897.5, 898.5, 899.5, 900.5, 901.5, 902.5, 903.5, 904.5, 905.5, 906.5, 907.5, 908.5, 909.5, 910.5, 911.5, 912.5, 913.5, 914.5, 915.5, 916.5, 917.5, 918.5, 919.5, 920.5, 921.5, 922.5, 923.5, 924.5, 925.5, 926.5, 927.5, 928.5, 929.5, 930.5, 931.5, 932.5, 933.5, 934.5, 935.5, 936.5, 937.5, 938.5, 939.5, 940.5, 941.5, 942.5, 943.5, 944.5, 945.5, 946.5, 947.5, 948.5, 949.5, 950.5, 951.5, 952.5, 953.5, 954.5, 955.5, 956.5, 957.5, 958.5, 959.5, 960.5, 961.5, 962.5, 963.5, 964.5, 965.5, 966.5, 967.5, 968.5, 969.5, 970.5, 971.5, 972.5, 973.5, 974.5, 975.5, 976.5, 977.5, 978.5, 979.5, 980.5, 981.5, 982.5, 983.5, 984.5, 985.5, 986.5, 987.5, 988.5, 989.5, 990.5, 991.5, 992.5, 993.5, 994.5, 995.5, 996.5, 997.5, 998.5, 999.5, 1000.5, 1001.5, 1002.5, 1003.5, 1004.5, 1005.5, 1006.5, 1007.5, 1008.5, 1009.5, 1010.5, 1011.5, 1012.5, 1013.5, 1014.5, 1015.5, 1016.5, 1017.5, 1018.5, 1019.5, 1020.5, 1021.5, 1022.5, 1023.5, 1024.5, 1025.5, 1026.5, 1027.5, 1028.5, 1029.5, 1030.5, 1031.5, 1032.5, 1033.5, 1034.5, 1035.5, 1036.5, 1037.5, 1038.5, 1039.5, 1040.5, 1041.5, 1042.5, 1043.5, 1044.5, 1045.5, 1046.5, 1047.5, 1048.5, 1049.5, 1050.5, 1051.5, 1052.5, 1053.5, 1054.5, 1055.5, 1056.5, 1057.5, 1058.5, 1059.5, 1060.5, 1061.5, 1062.5, 1063.5, 1064.5, 1065.5, 1066.5, 1067.5, 1068.5, 1069.5, 1070.5, 1071.5, 1072.5, 1073.5, 1074.5, 1075.5, 1076.5, 1077.5, 1078.5, 1079.5, 1080.5, 1081.5, 1082.5, 1083.5, 1084.5, 1085.5, 1086.5, 1087.5, 1088.5, 1089.5, 1090.5, 1091.5, 1092.5, 1093.5, 1094.5, 1095.5, 1096.5, 1097.5, 1098.5, 1099.5, 1100.5, 1101.5, 1102.5, 1103.5, 1104.5, 1105.5, 1106.5, 1107.5, 1108.5, 1109.5, 1110.5, 1111.5, 1112.5, 1113.5, 1114.5, 1115.5, 1116.5, 1117.5, 1118.5, 1119.5, 1120.5, 1121.5, 1122.5, 1123.5, 1124.5, 1125.5, 1126.5, 1127.5, 1128.5, 1129.5, 1130.5, 1131.5, 1132.5, 1133.5, 1134.5, 1135.5, 1136.5, 1137.5, 1138.5, 1139.5, 1140.5, 1141.5, 1142.5, 1143.5, 1144.5, 1145.5, 1146.5, 1147.5, 1148.5, 1149.5, 1150.5, 1151.5, 1152.5, 1153.5, 1154.5, 1155.5, 1156.5, 1157.5, 1158.5, 1159.5, 1160.5, 1161.5, 1162.5, 1163.5, 1164.5, 1165.5, 1166.5, 1167.5, 1168.5, 1169.5, 1170.5, 1171.5, 1172.5, 1173.5, 1174.5, 1175.5, 1176.5, 1177.5, 1178.5, 1179.5, 1180.5, 1181.5, 1182.5, 1183.5, 1184.5, 1185.5, 1186.5, 1187.5, 1188.5, 1189.5, 1190.5, 1191.5, 1192.5, 1193.5, 1194.5, 1195.5, 1196.5, 1197.5, 1198.5, 1199.5, 1200.5, 1201.5, 1202.5, 1203.5, 1204.5, 1205.5, 1206.5, 1207.5, 1208.5, 1209.5, 1210.5, 1211.5, 1212.5, 1213.5, 1214.5, 1215.5, 1216.5, 1217.5, 1218.5, 1219.5, 1220.5, 1221.5, 1222.5, 1223.5, 1224.5, 1225.5, 1226.5, 1227.5, 1228.5, 1229.5, 1230.5, 1231.5, 1232.5, 1233.5, 1234.5, 1235.5, 1236.5, 1237.5, 1238.5, 1239.5, 1240.5, 1241.5, 1242.5, 1243.5, 1244.5, 1245.5, 1246.5, 1247.5, 1248.5, 1249.5, 1250.5, 1251.5, 1252.5, 1253.5, 1254.5, 1255.5, 1256.5, 1257.5, 1258.5, 1259.5, 1260.5, 1261.5, 1262.5, 1263.5, 1264.5, 1265.5, 1266.5, 1267.5, 1268.5, 1269.5, 1270.5, 1271.5, 1272.5, 1273.5, 1274.5, 1275.5, 1276.5, 1277.5, 1278.5, 1279.5, 1280.5, 1281.5, 1282.5, 1283.5, 1284.5, 1285.5, 1286.5, 1287.5, 1288.5, 1289.5, 1290.5, 1291.5, 1292.5, 1293.5, 1294.5, 1295.5, 1296.5, 1297.5, 1298.5, 1299.5, 1300.5, 1301.5, 1302.5, 1303.5, 1304.5, 1305.5, 1306.5, 1307.5, 1308.5, 1309.5, 1310.5, 1311.5, 1312.5, 1313.5, 1314.5, 1315.5, 1316.5, 1317.5, 1318.5, 1319.5, 1320.5, 1321.5, 1322.5, 1323.5, 1324.5, 1325.5, 1326.5, 1327.5, 1328.5, 1329.5, 1330.5, 1331.5, 1332.5, 1333.5, 1334.5, 1335.5, 1336.5, 1337.5, 1338.5, 1339.5, 1340.5, 1341.5, 1342.5, 1343.5, 1344.5, 1345.5, 1346.5, 1347.5, 1348.5, 1349.5, 1350.5, 1351.5, 1352.5, 1353.5, 1354.5, 1355.5, 1356.5, 1357.5, 1358.5, 1359.5, 1360.5, 1361.5, 1362.5, 1363.5, 1364.5, 1365.5, 1366.5, 1367.5, 1368.5, 1369.5, 1370.5, 1371.5, 1372.5, 1373.5, 1374.5, 1375.5, 1376.5, 1377.5, 1378.5, 1379.5, 1380.5, 1381.5, 1382.5, 1383.5, 1384.5, 1385.5, 1386.5, 1387.5, 1388.5, 1389.5, 1390.5, 1391.5, 1392.5, 1393.5, 1394.5, 1395.5, 1396.5, 1397.5, 1398.5, 1399.5, 1400.5, 1401.5, 1402.5, 1403.5, 1404.5, 1405.5, 1406.5, 1407.5, 1408.5, 1409.5, 1410.5, 1411.5, 1412.5, 1413.5, 1414.5, 1415.5, 1416.5, 1417.5, 1418.5, 1419.5, 1420.5, 1421.5, 1422.5, 1423.5, 1

Db 299 TGCTCTTACAGCTGGGTGAGTTCACAGACGCCCTGAGCAGGCTCTCGCATGGCTGAC 358
 Qy 1322 ACACACCGAGGGCTTGTCTAATGACGAGAAACCTGTTGGAGAGACCTTAAAGCATTGA 1381
 Db 359 ACACACCGAGGGCTTGTCTAATGACGAGAAACCTGTTGGAGAGACCTTAAAGCATTGA 418
 Qy 1382 AATTGAACCTGCGAAGCATCATGCTCCAAATGATATTTAGCCCATGCTCAGCAGT 1441
 Db 419 GATTGAATGCTTAAGCATCATGCTCCAAACGACCTTTAGCCCATGCTCAGCAGT 478
 Qy 1442 GGAAGCCGTTAATAAGCAGGAAATGATCTAATTGAATCAAGTCAGAGAGAGAGAGCAG 1501
 Db 479 AGAAGCCGTTAATAAGCAGGAAATGATCTAAGTCAAGTCAGAGAGAGAGAGCAG 538
 Qy 1502 CAACCTTCAGAAACGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAAAAC 1561
 Db 539 CAACCTTCAGAAACGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAAAAC 598
 Qy 1562 AGAACAAGAGAGAGCAGCTGATGCTGCTGGCCAGGCAAAAGGTTCCATGGCGA 1621
 Db 599 AGATCAAGAGAGAGAGCAGCTGATGCTGCTGGCCAGGCAAAAGGTTCCATGGCGA 658
 Qy 1622 AATTGAAGATTGACAGCAGTGGCTGACTGACACGAGAGCGTATCTGTTGGCATCTAAAC 1681
 Db 659 AATGAGATTGGAACAGTGGCTGACGAGACAGAGCGTATCTGTTGGCATCTAAAC 718
 Qy 1682 GCTGGG 1687
 Db 719 TCTGGG 724

RESULT 12

CB723030 LOCUS 799 bp mRNA linear EST 09-JUL-2003
 DEFINITION UI-M-GH0-cel-a-06-0-UI.r1 NIH BMAP_GH0 Mus musculus cDNA clone
 CB723030 IMAGE:6840943 5', mRNA sequence.
 CB723030 VERSION 1
 CB723030.1 GI:29780172
 EST.
 Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 799)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@bbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mouse1.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES
 source
 Seq primer: PYX-5.
 Location/Qualifiers
 1..799
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone_image="6840943"
 /tissue_type="whole brain"
 /dev_stage="1, 5, and 15 days newborn"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_idb="NIH_BMAP_GH0"
 /note="Organ: Brain; Vector: PYX-Asc; Site 1: Ecor I;
 Site 2: Not I; The library was constructed according
 to Bontido, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Ecor I adaptor, digested with NotI and then cloned
 directionally into PYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CGAGCTGAAT. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'gene discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

Query Match 22.1%; Score 449.8; DB 6; Length 799;
 Best Local Similarity 76.5%; Pred. No. 4.2e-98;
 Matches 615; Conservative 0; Mismatches 177; Indels 12; Gaps 5;

Db 410 ATAAGATGTGTCTGATGACATGGAAGAGCTACAGCCGTGTATGAACCTCTTAAACGA 469
 Qy 1 ATAGAGCGTGTCCGTGACATGAGAAAGCTGACGCCGCTTACGACATCTGAGCGAGA 60
 Db 470 GGGGAGAGAAATGATTGCTAGATCTGGGGGAGCTGATAAGACATATCTGCCAAAGCTG 529
 Qy 61 GGGGAGAGAGATGATCGCCAGATCCGAGGGCACTGAAAAAGACGTATCTGCCAGAGCTG 120
 Db 530 TTCAGATTAAGCTTGACCAATGCTTTTCATTGGAAGAACATACACACTGCTGGAGAG 589
 Qy 121 TACAGATTAAGTTGAGCAAAATGTTTCATCTGGGAAAGCATCACACGCTGTGTGAAAG 180
 Db 590 AGAGGAGAGCCAACTACTGATGATGATGAGAGCTAGCAGAAAGTTCTGTGTGTATGACA 649
 Qy 181 AACGGAAAGCCAACTCTGTGATGATGAGAGCTAGCAGAAAGTTCTGTGTGTATGACA 240
 Db 650 TGTCAATTGATGATTACCATTAAGATCTCAAGATTTTCATCCGGGACCTGGAGATCTGTG 709
 Qy 241 TGTGCGTGTGTGTTACCATCAAGACACTCAAGATTTTCATCCGGGACCTGGAGATCTGTG 300
 Db 710 GAATTGATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 769
 Qy 301 GAATGACACCCCTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 Db 770 TAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 829
 Qy 361 TTGACGCGCTGACAGAGAGAGCTGACATGATGATGATGATGATGATGATGATGATGATGAT 420
 Db 830 CATGTGGGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 889
 Qy 421 CATGTGGGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
 Db 890 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 943
 Qy 480 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
 Db 944 GCAGCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1003
 Qy 540 GCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
 Db 1004 GGTAAAGTTGCTTCAATGATCTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1063
 Qy 600 AACAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 658
 Db 1064 TGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1122
 Qy 659 TGAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
 Db 1123 GATCTCATGAGCAGAGCTTTTGTGTAAGAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1182
 Qy 719 GAACATCATGAGCAGAGCTTTTGTGTAAGAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 775
 Db 1183 TTCTAGACCCAAATGAGACTGA 1206
 Qy 776 TCCAGAGACCGCTGATGAGACTGA 799

RESULT 13
BU217687 836 bp mRNA linear EST 25-NOV-2002
LOCUS 603759087F1 CSQCCHN04 Gallus gallus cDNA clone CHEST674n22 5', mRNA
DEFINITION
SEQUENCE.
ACCESSION BU217687.1 GI:25399501
VERSION BU217687.1
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE Boardman, P.E., Sanz-Bzquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
AUTHORS Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
1 (bases 1 to 836)
A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..836
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hixsex"
/db_xref="taxon:9031"
/clone="CHEST674n22"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_1ib="CSQCCHN04"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site_1: EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
1996) 791, except that a significantly longer
reannealing hybridization was used."

Query Match 21.7%; Score 441.6; DB 5; Length 836;
Best Local Similarity 76.7%; Pred. No. 4.2e-96;
Matches 540; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

ORIGIN
232 ATTAAGAAGATGTCAAAAAGCGTGTGCGACGTGATGAGCCATTCTCAATCACT 291
|||||
1 ATTAAGAAGATGTCAAAAAGAGACGACGTGCGATGATGAAGCCCTTCTCAGTGAAC 60
|||||
292 CAGTTCATGACGAAGATGATGATCCTTGAAGAGCTGGAAGCATCGTGAAGCTTGG 351
|||||
61 CAGTTCATGACGAAGATGATGATCCTTGAAGAGCTTGAAGCATGATGATGACCTGG 120
|||||
352 AGGACACCAACCTCTATCTCTGAGAGGTGAGAGATCAAGAGACAGATCACTGAAAT 411
|||||
Db 121 AGGACACCAACCTCTATCTCTGAGAGGTGAGAGATCAAGAGACAAATTTAGTGAAT 180
|||||
412 AAGATGTGTCACTAGACATGAAAAAGCTAACCCGTTGTATGAAACTTTAAACAGAG 471
|||||

Db 181 AAGATGTGTCAATGATCTGAAAAAAGCTTCCAGCCAGCTATGACGCTTAAAGCAGCA 240
|||||
472 GGAGAGGAATGATATGCTGATCTGGGGGAGCATGATTAAGACATATCTGCCAAAGCTTT 531
|||||
Db 241 GGAGAGGAATGATATGCTGATCTGGGGGAGCATGATTAAGACATATCTGCCAAAGCTTT 300
|||||
532 CAGATTAAGCTTGAACCAAAATGTTTTCATTTGGGAGACATACACACTGTGTGAAGAG 591
|||||
Db 301 CAAGATTAAGCTTGAACCAAAATGTTTTCATTTGGGAGACATACACACTGTGTGAAGAG 360
|||||
592 AGGAAAGCCAACTACTGATGTGATGAGCTTACGAGAAAGTTTGTGTGATCAGT 651
|||||
Db 361 AGGAGAGCCAAATGTTTGTGATGTATGAACTAGCGAAAGTTCTGGTGTGATCATAG 420
|||||
652 TCATGATATGATTAACATTAAGATCTCAAGATTCATCCGGGACCTGGAAGATCCTGGA 711
|||||
Db 421 GCTCTGTATGCTACTATTAAGATTAAGTACTCGAGCTTCATTCGAGAACTGAGAGGACCTGGA 480
|||||
712 ATTGATCTTCACTAGTAAACAAACAGACAGACAGACAGACCATTAAGGAGAAATA 771
|||||
Db 481 GTTGACCCACTGTGTATGATCAAGACAGACAGAGAGCTGAGGCTGCTAAAGAGAAAT 540
|||||
772 GATGACATACAGAGAGAGCTGATATGATTATTAACCTAGTTCGATCAATTCGGGCA 831
|||||
Db 541 GATGACATCAAGAGAGAGAGCTGATATGATTATTAACCTAGTTCGATCAATTCGGGCA 600
|||||
832 TGTGGGAGACCTGATTAACCAATTCGACAGAGAGATATCATGATTAATTCAGCGATG 891
|||||
Db 601 TGTGGAGACCTGACCAACCTATTTGTCACAGAGATATGATGATGATGATGATGATG 660
|||||
892 GGATTCCTTAATTAAGCTTGAAGAGACCGGATGACCACTTG 935
|||||
Db 661 GATGCTTAAACAAACATGAAACAGAGAGATGATTAAGCTTG 704
|||||

RESULT 14
BU368661 793 bp mRNA linear EST 28-NOV-2002
LOCUS 603597561F1 CSQCCHN73 Gallus gallus cDNA clone CHEST567115 5', mRNA
DEFINITION
SEQUENCE.
ACCESSION BU368661.1 GI:25876662
VERSION BU368661.1
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE Boardman, P.E., Sanz-Bzquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
AUTHORS Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
1 (bases 1 to 793)
A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..793
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton line 15T"
/db_xref="taxon:9031"
/clone="CHEST567115"
/sex="female"
/tissue_type="not cerebrum or cerebellum"


```

/dev stage="adult"
/lab_host="DH10B"
/clone_lib="CSECHN73"
/notes="Organ: Brain; Vector: Bluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dt) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
Bluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

```

ORIGIN

```

Query Match      21.7%; Score 440; DB 5; Length 793;
Best Local Similarity 75.4%; Pred. No. 1e-95;
Matches 576; Conservative 0; Mismatches 180; Indels 8; Gaps 2;

QY 244 GTCAAAAAGCGTGTGTCAGTGAAGCCATTCTCAATCACTCAGTTCATGAC 303
    |||||
DB 1 GTCAAAAAGAGCTGTGTCAGTGAAGCCCTTCTCAGTGAAGCCATTCTCAAT 60
    |||||
QY 304 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 363
    |||||
DB 61 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
    |||||
QY 364 TCTATCTCTGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
    |||||
DB 121 TCAATCTCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
    |||||
QY 424 GTAGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
    |||||
DB 181 ATGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
    |||||
QY 484 ATTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
    |||||
DB 241 ATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
    |||||
QY 544 GACCAATGCTTTCATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
    |||||
DB 301 GACCAATGCTTTCATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
    |||||
QY 604 CTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663
    |||||
DB 361 TTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
    |||||
QY 664 ACCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723
    |||||
DB 421 ACTATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
    |||||
QY 724 GTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 783
    |||||
DB 481 GTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
    |||||
QY 784 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
    |||||
DB 541 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
    |||||
QY 844 GATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902
    |||||
DB 601 GATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
    |||||
QY 903 -ATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
    |||||
DB 661 CAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
    |||||
QY 959 ---ACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 999
    |||||

```

DB 721 TACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 764

RESULT 15
LOCUS BQ215351 849 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT 7558465 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6046743
5' mRNA sequence.

ACCESSION BQ215351
VERSION BQ215351.1 GI:20396751
KEYWORDS EST.

SOURCE
ORGANISM Homo sapiens (human)

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 849)
NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

COMMENT
Contact: Robert Strauberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>

Plate: LLM13293 row: a column: 16
High quality sequence stop: 685.

FEATURES
Location/Qualifiers

1..849
/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6046743"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

```

Query Match      21.5%; Score 437.6; DB 5; Length 849;
Best Local Similarity 96.7%; Pred. No. 4e-95;
Matches 446; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1427 CCATCACTCCACAGTGAAGCCGTTAATTAAGCAGAAATGATTTAATGATCAAGTGC 1486
    |||||
DB 1 CCATCACTCCACAGTGAAGCCGTTAATTAAGCAGAAATGATTTAATGATCAAGTGC 60
    |||||
QY 1487 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1546
    |||||
DB 61 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
    |||||
QY 1547 TGTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1606
    |||||
DB 121 TGTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
    |||||
QY 1607 AGGTTTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1666
    |||||
DB 181 AGGTTTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
    |||||
QY 1667 GTTGGCATCTTAAACCGTGGAGAGGTTTACGGAGAGAGAGAGAGAGAGAGAGAGAG 1726
    |||||
DB 241 GTTGGCATCTTAAACCGTGGAGAGGTTTACGGAGAGAGAGAGAGAGAGAGAGAGAG 300
    |||||
QY 1727 TATGAGAGTGTGCTGCTCTTTGAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAG 1786
    |||||
DB 301 TATGAGAGTGTGCTGCTCTTTGAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAG 360
    |||||

```

Wed Jan 12 15:42:54 2005

us-10-089-887-46.rst

Page 12

OY	1787	AGGCACGAGATGCTGTGAAGAATGCCAAATTCGGAGAGCAAAATTTCACCAAGCAT	1846
Dδ	361	AGCCACGAGATGCTGTGCAATGCCAAATTCGGAGAGCAAAATTTCACCAAGCAT	420
OY	1847	AAATPACTTGAAAAAATAAAAAAAAAAAAAAAAAAAC	1887
Dδ	421	AAATPACTTGAAGAAAATGGGAATCGTGGAAACCAC	461

Search completed: January 12, 2005, 12:49:27
Job time : 5757 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2005, 05:49:47 ; Search time 925 Seconds
(without alignments)
11526.026 Million cell updates/sec

Title: US-10-089-887-46

Perfect score: 2031

Sequence: 1 ttgcacatgcgcacgaaga.....tngagagcccccacgcgc 2031

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N GeneSeq 23Sep04:*

1: geneSeq1980s:*

2: geneSeq1990s:*

3: geneSeq2000s:*

4: geneSeq2001as:*

5: geneSeq2001bs:*

6: geneSeq2002as:*

7: geneSeq2002bs:*

8: geneSeq2003as:*

9: geneSeq2003bs:*

10: geneSeq2003cs:*

11: geneSeq2003ds:*

12: geneSeq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1983	97.6	2028	6	ABK27463 DNA encod
2	1656.8	81.6	9274	6	ABK13224 DNA encod
3	1653.6	81.4	16684	12	ADQ24870 Human sof
4	1637.6	80.6	3307	11	ADM02768 Human cal
5	1010	49.7	6006	4	AAH41086 Human cal
6	655.4	32.3	6007	8	ACA56949 Human adi
7	655.4	32.3	16122	4	AAS22908 DNA encod
8	655.4	32.3	16341	4	AAS23002 DNA encod
9	501	24.7	2181	6	ABK34676 Human CDN
10	497.8	24.5	2113	4	AAS29811 Human cyt
11	497.4	24.5	5453	4	AAS29791 Human cyt
12	438.8	20.8	547	4	AAI22375 Human bre
13	422.8	20.8	531	4	AAI13506 Human bre
14	357	17.6	401	5	AAE65919 Novel hum
15	352.2	17.3	4327	4	AAK51484 Human pol
16	352.2	17.3	4438	4	AAK51483 Human pol
17	266.8	13.1	4495	4	AAI58467 Human pol
18	266.8	13.1	4495	5	ADQ98882 DNA encod
19	266.8	13.1	4495	9	ADB48442 Novel hum
20	266.8	13.1	4578	4	AAK52468 Human pol
21	266.8	13.1	4578	4	AAK52467 Human pol

ALIGNMENTS

22	266.8	13.1	4775	4	AAI60253 Human pol
23	262.8	12.9	344	6	ABN75109 Human str
24	250.2	12.3	278	3	AAC32058 Human sec
25	235.6	11.6	1721	6	ABBS1546 Human CDN
26	216.2	10.6	4721	5	AAAS4613 DNA encod
27	177.8	8.8	3684	10	ADA52704 Human cod
28	162.2	8.0	12995	4	ABL20575 Drosophi1
29	162.2	8.0	17678	4	ABL19341 Drosophi1
30	162.2	8.0	18011	4	ABL20689 Drosophi1
31	162.2	8.0	27845	4	ABL20591 Drosophi1
32	153.6	7.6	388	2	AAV90172 EST clone
33	135	6.6	39969	4	AAK85853 Human imm
34	134.6	6.6	563	12	ACH71003 Human gen
35	133.6	6.6	171	12	ACH84703 Human gen
36	115.8	5.7	31741	5	ABA21231 Human ner
37	93.8	4.6	1024	3	AAZ97148 Human pro
38	88.6	4.4	449	4	AAS29810 Human cyt
39	88.6	4.4	4156	4	AAS29790 Human cyt
40	87.4	4.3	22137	4	ABL20574 Drosophi1
41	87.4	4.3	48718	4	ABL20590 Drosophi1
42	87.4	4.3	63284	4	ABL20688 Drosophi1
43	87.4	4.3	71962	4	ABL19340 Drosophi1
44	82.6	4.1	3372	4	ABL1494 Human inh
45	69.4	3.4	1613	3	AAA07277 Human PST

RESULT 1
ID ABK27463 standard; CDNA, 2028 BP.
XX
AC ABK27463;
XX
DT 09-APR-2002 (first entry)
XX
DE DNA encoding colon cancer associated protein, CATX-14.
XX
KW Human; colon cancer; tumour; abnormal cell growth; melanoma;
KW cervical cancer; colorectal adenocarcinoma; Wilms' tumour; leukaemia;
KW lymphoma; antisense therapy; CATX; gene; ss.
XX
OS Homo sapiens.
XX
PN WO20011047-A2.
XX
PD 15-FEB-2001.
XX
PF 08-AUG-2000, 2000WO-US021606.
XX
PR 09-AUG-1999, 99US-0147933P.
XX
PA (FARB) BAYER CORP.
XX
PI Bowman BM, Wang K;
XX
XX WPI, 2002-121548/16.
DR P-PSDB; AAU81310.
XX
PT New isolated nucleic acid involved in growth regulation in human colonic
PT epithelial cells, termed CATX, for diagnosing and treating abnormal cell
PT growth, and for use as a probe/primer for detecting tumors.
XX
PS Claim 13; Page 122; 130pp; English.
XX
CC The invention relates to an isolated nucleic acid (I) involved in growth
CC regulation in human colonic epithelial cells, termed CATX. (I) is useful
CC as a probe/primer for detecting tumours, preferably colon cancer. The
CC nucleic acid, encoded polypeptide and antibody are useful in diagnosis
CC and treatment of abnormal cell growth (such as cervical cancer,
CC melanomas, colorectal adenocarcinomas, Wilms' tumour, leukaemia and
CC lymphomas), in screening assays for the treatment of abnormal cell

CC growth, for raising antibodies, and to screen for peptide analogues and
CC antagonists. (1) is useful as a biomarker for human tumor cells, e.g.,
CC colon cancer cells, for generating probes and primers designed for
CC identifying and/or cloning homologues in other cell types, in antisense
CC therapy, and in tissue profiling. (1) identifies cancer cells at an early
CC stage of development, so that premalignant cells can be identified prior
CC to their spreading throughout the human body. (1) allows the early detection
CC of potentially cancerous conditions, and treatment of the cancerous
CC conditions prior to spread of the cancer cells throughout the body, or
CC prior to development of an irreversible cancerous condition. ABK27426-
CC ABK27469 represent human colon cancer associated coding sequences and
CC primers of the invention. Note: This sequence is claimed as a polypeptide
CC in the specification

XX Sequence 2028 BP; 664 A; 412 C; 507 G; 436 T; 0 U; 9 Other;

Query Match 97.6%; Score 1983; DB 6; Length 2028;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2028; Conservative 0; Mismatches 2; Indels 5; Gaps 3;

QY 1 TTGCCATGCTGACAGAAACACAAATCAATCTCTCAGCTCCCGCCCGAGCCCTTGATA 60
DB 1 TTGCCATGCTGACAGAAACACAAATCAATCTCTCAGCTCCCGCCCGAGCCCTTGATA 60
QY 61 TGATCTCTAAGGACAGACAGAGAAACATCGCACTGCGTAGTTGATGACTGACACA 120
DB 61 TGATCTCTAAGGACAGACAGAGAAACATCGCACTGCGTAGTTGATGACTGACACA 120
QY 121 CAAGCCCTCATATGATGATGATGAAACAAACTGGCCACAGTACGATGAGCCCTGGGAA 180
DB 121 CAAGCCCTCATATGATGATGATGAAACAAACTGGCCACAGTACGATGAGCCCTGGGAA 180
QY 181 GGCTTTCTATCCAGAGAAAGTAGTGTGGACGCGACACCTTTACAGTCACTTAAAGAA 240
DB 181 GGCTTTCTATCCAGAGAAAGTAGTGTGGACGCGACACCTTTACAGTCACTTAAAGAA 240
QY 241 GATGTCAAAAAAGCGTGTGTGGACATGATGAAAGCCATTTCTCAATCACTCAGTTCCAT 300
DB 241 GATGTCAAAAAAGCGTGTGTGGACATGATGAAAGCCATTTCTCAATCACTCAGTTCCAT 300
QY 301 GAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 301 GAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 CCTCTATCTCTCAGAGGTTGAGAGATCAAGAAACAGATCAGTGAATAAAGAAATG 420
DB 361 CCTCTATCTCTCAGAGGTTGAGAGATCAAGAAACAGATCAGTGAATAAAGAAATG 420
QY 421 TCAGTAGACATGGAAGAGCTACAGCGTGTATGAACTTTAAACAGAGGGAGAGAA 480
DB 421 TCAGTAGACATGGAAGAGCTACAGCGTGTATGAACTTTAAACAGAGGGAGAGAA 480
QY 481 ATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 ATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 CTTCGACCAATGTTTTTCAATTTGGAGAACATACACACTGTTGAGAGAGAGAGGCC 600
DB 541 CTTCGACCAATGTTTTTCAATTTGGAGAACATACACACTGTTGAGAGAGAGAGGCC 600
QY 601 AAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 AAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GTTACATTTAAAGATCTCAAGTTTCTATCCGGGACCTGGAGATCTGGAAATTTGATCCT 720
DB 661 GTTACATTTAAAGATCTCAAGTTTCTATCCGGGACCTGGAGATCTGGAAATTTGATCCT 720
QY 721 TCAGTAGTAAACAAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 TCAGTAGTAAACAAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 CAGGAGAGAGCTGATATGATTAATTAACCTAGGTTCTGAACTCATTCGCGGAGTGGGAG 840

DB 781 CAGGAGAGAGCTGATATGATTAATTAACCTAGGTTCTGAACTCATTCGCGGAGTGGGAG 840
QY 841 CCTGATTAACCCATTTGTCAGAAAGATATATGATTAATTTACAGGATGGATTTCTCT 900
DB 841 CCTGATTAACCCATTTGTCAGAAAGATATATGATTAATTTACAGGATGGATTTCTCT 900
QY 901 AATTAAGCTTGAAGAACCGGATTTGACCACTTGAGAGCAATGACGTCGCTGATAC 960
DB 901 AATTAAGCTTGAAGAACCGGATTTGACCACTTGAGAGCAATGACGTCGCTGATAC 960
QY 961 AGATGACTGACAGCGGATATTTTCTGGGTAGATTAATGCAAGTGTGATTTGCTTCAAT 1020
DB 961 AGATGACTGACAGCGGATATTTTCTGGGTAGATTAATGCAAGTGTGATTTGCTTCAAT 1020
QY 1021 GTCTCCAAATTTGAACAGCATCTTCGAAACCTGTCAGACAGAGATGATGTCTAGACAT 1080
DB 1021 GTCTCCAAATTTGAACAGCATCTTCGAAACCTGTCAGACAGAGATGATGTCTAGACAT 1080
QY 1081 TTAAGTCTGAGGCTTATCAACAGCAGATGAGTGTACGACTGACTCATCAGCAGACGC 1140
DB 1081 TTAAGTCTGAGGCTTATCAACAGCAGATGAGTGTACGACTGACTCATCAGCAGACGC 1140
QY 1141 TTTTGCTAAAGAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 TTTTGCTAAAGAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 ANCTGATTTGATGAT - GGTAGCCTGTGAGAGAGAGATCATCAACAGACATTAAC 1259
DB 1201 ANCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
QY 1260 AGGGTCTCTATTAAGCCNTGGGTCAAGTTCCAAACATCCCTGTATGAGCTCTCGACGCTG 1319
DB 1260 AGGGTCTCTATTAAGCCNTGGGTCAAGTTCCAAACATCCCTGTATGAGCTCTCGACGCTG 1319
QY 1320 ACAACACCGAGGCTTGTGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1379
DB 1320 ACAACACCGAGGCTTGTGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1379
QY 1380 GAAATTTGAACCTTGCCAGACATATGATGCTCCAAATATGATGATTTAGCCATCAGTCA 1439
DB 1380 GAAATTTGAACCTTGCCAGACATATGATGCTCCAAATATGATGATTTAGCCATCAGTCA 1439
QY 1440 GTGGAAGCCGTTAATTAAGCAGAGAAATGATCTTAATTAAGTCAAGTCAAGAGAGAGAG 1499
DB 1440 GTGGAAGCCGTTAATTAAGCAGAGAAATGATCTTAATTAAGTCAAGTCAAGAGAGAGAG 1499
QY 1497 AGCAACCTTCAGAACAAAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAAA 1556
DB 1497 AGCAACCTTCAGAACAAAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAAA 1556
QY 1560 ACAGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1619
DB 1560 ACAGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1619
QY 1620 GAAATTTGAGGATTTGACAGAGTGTGCTGACCTGACAGAGAGAGAGAGAGAGAGAGAG 1679
DB 1620 GAAATTTGAGGATTTGACAGAGTGTGCTGACCTGACAGAGAGAGAGAGAGAGAGAGAG 1679
QY 1679 CCGCTGGAGGTTTACCGGAAACAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1739
DB 1679 CCGCTGGAGGTTTACCGGAAACAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1739
QY 1740 GCTGCTTTGAAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1799
DB 1740 GCTGCTTTGAAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1799
QY 1800 CTTGCAAGATGCGCAAAATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1859
DB 1800 CTTGCAAGATGCGCAAAATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1859
QY 1859 CTTGCAAGATGCGCAAAATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1919
DB 1859 CTTGCAAGATGCGCAAAATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1919

Db 1857 AAAAAAAAAAAAAAAAAAAAAAAAAANACGCCGCTGATTTAGTAATTCAGCGCCGCT 1916
Oy 1920 GAATTTAGACCTGCGCGCGCGCGCTCGAGCCCTTACTGAGTCTTATTAGATGGA 1979
Db 1917 GAATTTAGACCTGCGCGCGCGCGCTCGAGCCCTTACTGAGTCTTATTAGATGGA 1976
Oy 1980 ATCACTAGTCCGCGCGCTGAGTGCACCATATNGAGAGCTCCCAAGCGCT 2031
Db 1977 ATCACTAGTCCGCGCGCGCTGAGTGCACCATATNGAGAGCTCCCAAGCGCT 2028

RESULT 2
ABK13224
ID ABK13224 standard; cDNA; 9274 BP.
XX
AC ABK13224;
XX
DT 09-APR-2002 (first entry)
XX
DE DNA encoding novel signal pathway protein protein PN9109.
XX
KM MAPKAP-K3, Ap-3 delta; AP-695; amyloid A4; Hep8; heat shock protein 8;
KM L130; NY-REN-58; P38 Alpha; P38 Beta; ERK3; KIAA0934; CDR9;
KM cell cycle dependent kinase 9; CLK; C-NAP-1; clathrin heavy chain;
KM amphipysin; PN9109; KIAA1106; neurodegenerative disorder;
KM non-insulin dependent diabetes mellitus; NIDDM; Alzheimer's disease;
KM inflammatory disease; rheumatoid arthritis; inflammatory bowel disorder;
KM atherosclerosis; cardiac hypertrophy; hypoxic brain injury;
KM yeast two-hybrid; signal transduction pathway; human; ss;
KM mlogn activated protein kinase.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..8508
FT /*tag= a
FT /product= "PN9109"
FT /partial
FT /note= "No start codon given"
XX
PN WO200198524-A2.
XX
PD 27-DEC-2001.
XX
PF 21-JUN-2001; 2001WO-US019762.
XX
PR 22-JUN-2000; 2000US-0213245P.
XX
PA (MYRI-) MYRIAD GENETICS INC.
XX
PI Heichman K, Bartel PL;
XX
DR WP1; 2002-122287/16.
DR P-PSDB; AAU75100.
XX
PT New protein complexes comprising protein-protein interactions (e.g.
PT MAPKAP-K3/AP-3 delta or C-NAP-1/Clathrin HC), useful for diagnosing
PT physiological generative disorders or screening drugs for these diseases.
XX
PS Example 12; Page 27-31; 60pp; English.
XX
CC The invention describes an isolated protein complex, comprising two
CC proteins. The protein complex comprises: protein kinase MAPKAP-K3 and AP-
CC 3 adaptor complex delta sub-unit; MAPKAP-K3 and amyloid A4 precursor
CC protein, AP-695; MAPKAP-K3 and heat shock protein (Hsp) 8; leucine rich
CC L130 and NY-REN-58; P38 Alpha and P38 Beta; protein kinase ERK3 and
CC KIAA0934 (unknown function); ERK3 and cell cycle dependent kinase (CDK) 9
CC ; ERK3 and protein kinase CLK; C-NAP-1 and Clathrin heavy chain; C-NAP-1
CC and Amphipysin; C-NAP-1 and novel protein PN9109 or C-NAP-1 and KIAA1106
CC (unknown function) interactions. The protein complexes are useful for
CC diagnosing physiological generative disorders, drug screening for agents
CC that modulate the interaction of the proteins (thus identify drug
CC targets), and identifying additional proteins in the pathway common to

CC the proteins. These physiological disorders include non-insulin dependent
CC diabetes mellitus (NIDDM), neurodegenerative disorders (e.g. Alzheimer's
CC disease), inflammatory diseases (e.g. rheumatoid arthritis and
CC inflammatory bowel disorder) and other human disease such as
CC atherosclerosis, cardiac hypertrophy and hypoxic brain injury. This
CC sequence encodes the novel protein PN9109, found to bind C-NAP1 in a
CC yeast two-hybrid assays used to determine components of signal
CC transduction pathways, described in the method of the invention
XX
SQ Sequence 9274 BP; 3055 A; 1878 C; 2241 G; 2100 T; 0 U; 0 Other;
Query Match 81.6%; Score 1656.8; DB 6; Length 9274;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1805; Conservative 0; Mismatches 72; Indels 18; Gaps 9;
Oy 8 GCTGACAGAAACACATCATCTCTCAGC-TCGCCGCCAGCCCTTGAAATGATAC 66
Db 3612 GCTGACAGAAACACATCATCTCTCAGCCTTCCGCCCCAGCCCTTGAAATGAAAC 3671
Oy 67 TCTAAGCAGCAGAGAGAAACATCGGCACTGCGTGAAGTTAGCTGAACACAGACC 126
Db 3672 TCTAAGCAGCAGAGAGAAACATCGGCACTGCGTGAAGTTAGCTGAACACAGACC 3721
Oy 127 TCATATAGATTAAGTGAACAAACTGG----CCACAGTACTGATGAGCCCTGGGGAAG 182
Db 3732 TCATATAGATTAAGTGAACAAACTGGGCCACAGTTACTGAAATTGAGCCCTGGGGAAG 3791
Oy 183 CTTTTCATCCAGAGAGATGTGTGGACGCCGACCCCTTTACAGTCACATTAAGAAGA 242
Db 3792 CTTTTCATCCAGAGAGATGTGTGGACGCCGACCCCTTTACAGTCACATTAAGAAGA 3851
Oy 243 TGTCAAAAAGGTCTGTGGACCTGAGATGAAGCCATTTCTCAATCACTCAGTTCCATGA 302
Db 3852 TGTCAAAAAGGTCTGTGGACCTGAGATGAAGCCATTTCTCAATCACTCAGTTCCATGA 3911
Oy 303 CAAGATGATCAGATCTCTTGAGACCTGGAACGATGCTGGAAGCTTGAAGCAGCCACC 362
Db 3912 CAAGATGATCAGATCTCTTGAGACCTGGAACGATGCTGGAAGCTTGAAGCAGCCACC 3971
Oy 363 CTCTATCTCTGACAGAGTTGAGAAATCAAGAAACAGATCGTAAGAAATTAAGTGTCTC 422
Db 3972 CTCTATCTCTGACAGAGTTGAGAAATCAAGAAACAGATCGTAAGAAATTAAGTGTCTC 4031
Oy 423 AGTAGACATGGAAGGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGAGAGAAAT 482
Db 4032 AGTAGACATGGAAGGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGAGAGAAAT 4091
Oy 483 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAGCTGTTCAAGATTAAGCT 542
Db 4092 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAGCTGTTCAAGATTAAGCT 4151
Oy 543 TGACCAATGGTTTTCATTTGGGAGAACATACACACTGGTGGAGAGAGAGCCAA 602
Db 4152 TGACCAATGGTTTTCATTTGGGAGAACATACACACTGGTGGAGAGAGAGCCAA 4211
Oy 603 ACTACTGATGTATGAGAGCTAGACGAAAGTTCTGTGTGATCACTGTCATTAAGTACT 662
Db 4212 ACTACTGATGTATGAGAGCTAGACGAAAGTTCTGTGTGATCACTGTCATTAAGTACT 4271
Oy 663 TACCATTAAGATCTCAAGATTTTCATCCGGGAGCTGGAAGATCTGGAATTAAGCTTTC 722
Db 4272 TACCATTAAGATCTCAAGATTTTCATCCGGGAGCTGGAAGATCTGGAATTAAGCTTTC 4331
Oy 723 AGTAGTAACCAACAGCAGAGAGAGACCATTAAGGAGAGAAATAGATGAGACTACA 782
Db 4332 AGTAGTAACCAACAGCAGAGAGAGACCATTAAGGAGAGAAATAGATGAGACTACA 4391
Oy 783 GGAGAGAGCTGATTAAGTTATTAACCTAGAGTTCTGAACCTATTCGGCAGTGGGAGCC 842
Db 4392 GGAGAGAGCTGATTAAGTTATTAACCTAGAGTTCTGAACCTATTCGGCAGTGGGAGCC 4451
Oy 843 TGATTAACCCATTTGCAAGAGAGATATACATGATTAATTCAGCGATGGGATCTCTAA 902

QY 303 CAAGATGATGATGATCCCTTGAGAGCTGGAAAGCATGCTGGAAGCTCTGAGAGCCAGCCACC 362
DB 11032 CAAGATGATGATGATCCCTTGAGAGCTGGAAAGCATGCTGGAAGCTCTGAGAGCCAGCCACC 11091
QY 363 CTCTATCTCTGACAGAGTTGAGAAAGATCAAGAAACAGATCAAGTAAATAAGATGTC 422
DB 11092 CTCTATCTCTGACAGAGTTGAGAAAGATCAAGAAACAGATCAAGTAAATAAGATGTC 11151
QY 423 AGTAGACATGAGAAAGCTACAGCCGTTGATGAAACTTTAACAAGAGGAGAGAAAT 482
DB 11152 AGTAGACATGAGAAAGCTACAGCCGTTGATGAAACTTTAACAAGAGGAGAGAAAT 11211
QY 483 GATTCGTAAGATCTGCGGGGAGCTGATTAAGACATATCTGCCAACTGCTTCAGTAAAGT 542
DB 11212 GATTCGTAAGATCTGCGGGGAGCTGATTAAGACATATCTGCCAACTGCTTCAGTAAAGT 11271
QY 543 TGACCAATGAGTTTCTATTGGGAGAAATATACACACTGCTGGAAGAGAGAGCCAA 602
DB 11272 TGACCAATGAGTTTCTATTGGGAGAAATATACACACTGCTGGAAGAGAGAGCCAA 11331
QY 603 ACTACGTGATGTGATGAGCTAGCAGAAAAGTTCTGCTGATGATCAATGTCATTGATAGT 662
DB 11332 ACTACGTGATGTGATGAGCTAGCAGAAAAGTTCTGCTGATGATCAATGTCATTGATAGT 11391
QY 663 TACCATTAAGATATCTCAAGTTTCAATCCGGAGCTGGAAGATCTCGGAATGATCTCTTC 722
DB 11392 TACCATTAAGATATCTCAAGTTTCAATCCGGAGCTGGAAGATCTCGGAATGATCTCTTC 11451
QY 723 AGTAGTAAACCAACAGCAGAAAGAGCAGAGACCATTAAGGAGAAATATAGTGAATCA 782
DB 11452 AGTAGTAAACCAACAGCAGAAAGAGCAGAGACCATTAAGGAGAAATATAGTGAATCA 11511
QY 783 GGAGAGCTGAGATATGATTATTAACCTAGTTCTGAATCTCATTCGCGATGTGGAGACC 842
DB 11512 GGAGAGCTGAGATATGATTATTAACCTAGTTCTGAATCTCATTCGCGATGTGGAGACC 11571
QY 843 TGATTAACCCATTTGTCAGAAAGATATCAAGATTAATTCAGCGATGGATCTCTTA 902
DB 11572 TGATTAACCCATTTGTCAGAAAGATATCAAGATTAATTCAGCGATGGATCTCTTA 11630
QY 903 ATAAAGCTTGGAAAGACCGAGTTGACCACTTGAG-----AGCAATGCAGCTGCCGTCA 956
DB 11631 ATAAAGCTTGGAAAGACCGAGTTGACCACTTGAGAGGCAATGCAGCTGCCGTCA 11690
QY 957 GTACAGATGACCTGACGCGGTATTTTCTGCTGATATTTGCAAGTGTGTAAGTTCGCTT 1016
DB 11691 ACCAGATGACCTGACGCGGTATTTTCTGCTGATATTTGCAAGTGTGTAAGTTCGCTT 11750
QY 1017 CAATGCTCTCCAAATTGGAACAGCATCTGAAACTGTCAAGCAGCATGATGCTGTAAG 1076
DB 11751 CAATGCTCTCCAAATTGGAACAG-ATCTGGAACCTGTCAAGCAGCATGATGGAAGCTTAAG 11809
QY 1077 CAATTAAGTGTGAGGCTATCAACAGCAGTAG-AGTGAAGCACTGACATCAAGCA 1135
DB 11870 GA-CTTTTGTGTAAGAAAGTAACAGAAAGAGTGAACAAACACTGTTCAAGACCATT 11928
QY 1136 GACGCTTTTGTGTAAGAAAGTAACAGNAGAGTGAACAAACACTGTTCTGAGCCCAA 1195
DB 11870 GA-CTTTTGTGTAAGAAAGTAACAGAAAGAGTGAACAAACACTGTTCAAGACCATT 11928
QY 1196 AATGAGNCTGATATTTGATAT-GGCTAGCTGTGAGNAGATCATCAACAGCAGCATTA 1254
DB 11929 AATGAGNCTGATATTTGATATTTGATGAGTGTGAGNAGATCATCAACAGCAGCATTA 11988
QY 1255 ACTGAGAGGATGCTCATATTAGCCMTGGGTCAGTTCCAAATGAGCCCTGTATGAGCTCTG- 1312
DB 11989 ACTGAGAGGATGCTCATATTAGCCMTGGGTCAGTTCCAAATGAGCCCTGTATGAGCTCTG- 12048
QY 1313 CATGCTGACACACACCGAGGGCTTGTAAAGTGAAGAGAAACCTGTGAGAGAGACCTTA 1372
DB 12049 ATGCTGACACACACCGAGGGCTTGTAAAGTGAAGAGAAACCTGTGAGAGAGACCTTA 12108

QY 1373 AGCATTAAGAAATTGAACCTTGACAAAGCATATGTCCTCCAAAATGATGATTAAGCCATCA 1432
DB 12109 AGCATTAAGAAATTGAACCTTGACAAAGCATATGTCCTCCAAAATGATGATTAAGCCATCA 12168
QY 1433 GTCCACAGTGAAGCCCTTAATTAAGCAGGAAATGATCTAATTAATCAAGTGCAGAGAGA 1492
DB 12169 GTCCACAGTGAAGCCCTTAATTAAGCAGGAAATGATCTAATTAATCAAGTGCAGAGAGA 12228
QY 1493 AGAAGCAAGCACTTTCAGAAACAAAGCTAGAGCTTTAAATCAACGCTGCAAAATGTTTT 1552
DB 12229 AGAAGCAAGCACTTTCAGAAACAAAGCTAGAGCTTTAAATCAACGCTGCAAAATGTTTT 12288
QY 1553 GGAAGAAACAGAAAGAAAGAGAGCAGCTGATGATGTCCTTGGCCAGGCAAGGATT 1612
DB 12289 GGAAGAAACAGAAAGAAAGAGAGCAGCTGATGATGTCCTTGGCCAGGCAAGGATT 12348
QY 1613 CCATGGCAAAATTGAGATTTGACAGAGTGGCTGACTGACACGAGGCTCATCTGTTGGC 1672
DB 12349 CCATGGCAAAATTGAGATTTGACAGAGTGGCTGACTGACACGAGGCTCATCTGTTGGC 12408
QY 1673 ATCTAAACCGCTGGAGAGTTTACCGAAACAGCCAGAGAGCAGCTTAATGCTCATATGA 1732
DB 12409 ATCTAAACCGCTGGAGAGTTTACCGAAACAGCCAGAGAGCAGCTTAATGCTCATATGA 12468
QY 1733 AGTCTGTCGCTTGTGAAGCTAAGAAAGAAACATATAGAGTGTGATGACAGAAAGGCA 1792
DB 12469 AGTCTGTCGCTTGTGAAGCTAAGAAAGAAACATATAGAGTGTGATGACAGAAAGGCA 12528
QY 1793 GCAGATGCTTGCAGAGATGCCCAAAATCTGCAGAGACAAATATTTGACCAAGACATTAATA 1852
DB 12529 GCAGATGCTTGCAGAGATGCCCAAAATCTGCAGAGACAAATATTTGACCAAGACATTAATA 12588
QY 1853 CTTGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1887
DB 12589 CTTGAAAAAATAATGGGAATCGGTGAAACCAAC 12623

RESULT 4
ADM02768
ID ADM02768 standard; cDNA: 3307 BP.
XX AC
XX ADM02768;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human cDNA of the invention SEQ ID NO:1453.
XX
XX seq; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
XX Homo sapiens.
XX
XX BP1347046-A1.
XX
XX 24-SEP-2003.
XX
XX 12-APR-2002; 2002EP-00008400.
XX
XX 22-MAR-2002; 2002JP-00137785.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Iobgai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX WPI: 2003-723558/69.
XX P-FSDB: ADM05211.
XX
XX New polynucleotides and polypeptides are useful in gene therapy, for
XX developing a diagnostic marker or medicines for regulating their
XX expression and activity, or as a target of gene therapy.
XX
XX Claim 1; SEQ ID NO 1453; 305pp; English.

XX The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM0316-ADM0358 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC cDNA sequence of the invention.

SQ Sequence 3307 BP; 1122 A; 632 C; 829 G; 724 T; 0 U; 0 Other;

Query Match 80.6%; Score 1637.6; DB 11; Length 3307;
 Best Local Similarity 95.9%; Pred. No. 0;
 Matches 1778; Conservative 0; Mismatches 58; Indels 18; Gaps 9;

QY 8 GCTGACAGAAACAAATCAATCTCTCTAGC-TCCCGCCCCAGCCCTTGAATATGATAC 66
 DB 1457 GCTGACAGAAACAAATCAATCTCTCTAGCCTTCCCGCCCCAGCCCTTGAATATGAAAC 1516
 QY 67 TCTAAGCCAGCAGCAGAGAGAAATCGGCACTGCGTGAATTATAGCTGAACAAGCC 126
 DB 1517 TCTAAGCCAGCAGCAGAGAGAAATCGGCACTGCGTGAATTATAGCTGAACAAGCC 1576
 QY 127 TCATATAGATTAAGATGAACAAAAGTGG-CCACAGTACTGAATGAGCCCTGGGGAAG 182
 DB 1577 TCATATAGATTAAGATGAACAAAAGTGGCCACAGTACTGAATGAGCCCTGGGGAAG 1636
 QY 183 CTTTCTATCCAGAGAGAGATATGTGCAAGCCGACACCTTTACAGTCACTTAAGAAGA 242
 DB 1637 CTTTCTATCCAGAGAGAGATATGTGCAAGCCGACACCTTTACAGTCACTTAAGAAGA 1696
 QY 243 TGTCAAAAAGCGTGTGTGCACTGATGAAGCATTTTCAATCAACTGATTTCCATGA 302
 DB 1697 TGTCAAAAAGCGTGTGTGCACTGATGAAGCATTTTCAATCAACTGATTTCCATGA 1756
 QY 303 CAAGATAGATCAGATCCTTGAGAGCCTGGAAGCATGTGGAAGCTGTGAGGAGCCACC 362
 DB 1757 CAAGATAGATCAGATCCTTGAGAGCCTGGAAGCATGTGGAAGCTGTGAGGAGCCACC 1816
 QY 363 CTCTATCTCTGCAAGGTTGAGAAGATCAAGAAAGATCAAGTGAATTAAGATGTGTC 422
 DB 1817 CTCTATCTCTGCAAGGTTGAGAAGATCAAGAAAGATCAAGTGAATTAAGATGTGTC 1876
 QY 423 AGTAGACATGGAAGAAAGCTACAGCCGTTGTATGAACCTTTAAACAAGGGGAGAGAAAT 482
 DB 1877 AGTAGACATGGAAGAAAGCTACAGCCGTTGTATGAACCTTTAAACAAGGGGAGAGAAAT 1936
 QY 483 GATTGCTAGATCTGGGGGGAAGCTGATTAAGACATATGCGCAAGCTGTTCAAGATTAAGCT 542
 DB 1937 GATTGCTAGATCTGGGGGGAAGCTGATTAAGACATATGCGCAAGCTGTTCAAGATTAAGCT 1996
 QY 543 TGAACAAATGTTTTTCAATTTGGAGAACATACACACTGCTGTGGAAGAGAGAGAGCCAA 602
 DB 1997 TGAACAAATGTTTTTCAATTTGGAGAACATACACACTGCTGTGGAAGAGAGAGAGCCAA 2056
 QY 603 ACTACTGATGTATGAGAGTACAGAAAGTTCTGCTGTGATCACTGTCATTGATAGT 662
 DB 2057 ACTACTGATGTATGAGAGTACAGAAAGTTCTGCTGTGATCACTGTCATTGATAGT 2116
 QY 663 TACCATTAAGATTAAGATCAAGATTTCACTCGGGAAGCTGGAAGATCTGGAATTAAGCTTTC 722
 DB 2117 TACCATTAAGATTAAGATCAAGATTTCACTCGGGAAGCTGGAAGATCTGGAATTAAGCTTTC 2176
 QY 723 AGTAGTAAACCAACAGCAAGAACAGCAAGACCATTAAGGGAAGAAATAGATGACTACA 782
 DB 2177 AGTAGTAAACCAACAGCAAGAACAGCAAGACCATTAAGGGAAGAAATAGATGACTACA 2236
 QY 783 GGAAGAGCTGAGATATAGTTTAACTAGGTTTGAATCTGATGCGGACATGTGGGAGGCC 842
 DB 2237 GGAAGAGCTGAGATATAGTTTAACTAGGTTTGAATCTGATGCGGACATGTGGGAGGCC 2296

QY 843 TGATAAACCCATTTGCAAGAGATATACATGATTAATTCAGCATGGATTTCTCTAA 902
 DB 2297 TGATAAACCCATTTGCAAGAGATATACATGATTAATTCAGCATGGATTTCTCTAA 2355
 QY 903 ATTAAGCTTGAAGAGCCGATTTGACCACTTGAG-----AGCAATGACGTCGCGTCA 956
 DB 2356 ATTAAGCTTGAAGAGCCGATTTGACCACTTGAGAGGCAATGACGCTGCGGTTCACT 2415
 QY 957 GTRACATGATGATGCAAGGCGGATTTTTTCTGGGTATATATGCAAGTGTGAAGTGTCTT 1016
 DB 2416 ACCAGATGATGATGCAAGGCGGATTTTTTCTGGGTATATATGCAAGTGTGAAGTGTCTT 2475
 QY 1017 CAATGCTCCAAATTTGGAACAGCATCTCGAAACCTGTCAACAGAGATTTGATGCTAGAG 1076
 DB 2476 CAATGCTCCAAATTTGGAACAGCATCTCGAAACCTGTCAACAGAGATTTGATGCTAGAG 2534
 QY 1077 CAATTTAAGCTGAGGCTTATCAACAGAGATAG-AGTGTAGCACTGACTATCAAGCA 1135
 DB 2535 CAATTTAAGCTGAGGCTTATCAACAGAGATAGAGAAATGGAAGATGATCATCAAGCA 2594
 QY 1136 GACGCTTTGCTTAAAGAAAGTAAACAGAGAGATGCAAAACACTGTTCTAGACCCAA 1195
 DB 2595 GA-GCTTTTGTAAAGAAAGTAAACAGAGAGATGCAAAACACTGTTCTAGACCCAA 2653
 QY 1196 AATGANTGATATGATATAT-GGCTAGCCTGAGAGNAGAGATCATCAACAGAGATTA 1254
 DB 2654 AATGANTGATATGATATATGAGATAGCTTGAAGAGAGAAATCATCAACAGAGATTA 2713
 QY 1255 ACTGAGAGGTGCTATTAAGCCNTGGGTGAGTTCACATGCGCTGATAGGCTCTG-- 1312
 DB 2714 ACTGAGAGGTGCTATTAAGCCNTGGGTGAGTTCACATGCGCTGATAGGCTCTG-- 2773
 QY 1313 CATGCTGACACACACGAGAGGCTTGTAGTAGAGCAAAACCTGTTGAGAGAGACCTTA 1372
 DB 2774 ATGGCTGACACACACGAGAGGCTTGTAGTAGAGCAAAACCTGTTGAGAGAGACCTTA 2833
 QY 1373 AGCCATTAAGATTTGAACCTTGCAAGCATATGCTGCCAAATATGATATTAAGCCATCA 1432
 DB 2834 AGCCATTAAGATTTGAACCTTGCAAGCATATGCTGCCAAATATGATATTAAGCCATCA 2893
 QY 1433 GTTCCACATGGAAGCCGTTAATTAAGCAGAGAAATGATCTTAATTAATCAAGTGCAGAGA 1492
 DB 2894 GTTCCACATGGAAGCCGTTAATTAAGCAGAGAAATGATCTTAATTAATCAAGTGCAGAGA 2953
 QY 1493 AGAAGCAACCACTTCAAGACAGTGAAGTTTAAATCAACGCTGGCAAAATGTTTT 1552
 DB 2954 AGAAGCAACCACTTCAAGACAGTGAAGTTTAAATCAACGCTGGCAAAATGTTTT 3013
 QY 1553 GGAAGAAACAGAACAAAGAGACAGCAGCTGATGCTGCTGGCCAGGCCAAAGGTT 1612
 DB 3014 GGAAGAAACAGAACAAAGAGACAGCAGCTGATGCTGCTGGCCAGGCCAAAGGTT 3073
 QY 1613 CCATGGCGAAATTTGAGATTTTGCAGCAGTGTGCTGA-CTGACAGAGCGTCATCTGTTGCG 1672
 DB 3074 CCATGGCGAAATTTGAGATTTTGCAGCAGTGTGCTGA-CTGACAGAGCGTCATCTGTTGCG 3133
 QY 1673 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGAGCTTAATGTCCATTTGGA 1732
 DB 3134 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGAGAGCTTAATGTCCATTTGGA 3193
 QY 1733 AGTGTGTGCTGCTTTGAAGCTTAAAGAAACATTAAGAGCTGATGACAGAAAGGCA 1792
 DB 3194 AGTGTGTGCTGCTTTGAAGCTTAAAGAAACATTAAGAGCTGATGACAGAAAGGCA 3253
 QY 1793 GCAGATGCTTGCAGAGATGCCAAATCTGCAGAGACAAATATTTGACCAAGACAT 1846
 DB 3254 GCAGATGCTTGCAGAGATGCCAAATCTGCAGAGACAAATATTTGACCAAGACAT 3307

RESULT 5
 AAH41086
 ID AAH41086 standard; cDNA; 6006 BP.

ID	ACAA56949 standard; cDNA; 6007 BP.
XX	AC ACAA56949;
XX	DT 10-JUN-2003 (first entry)
XX	DE Human adipocyte Selected Interacting domain, SID, cDNA #36.
XX	KW Human; ss; gene; prey; adipocyte; SID; selected interacting domain;
XX	KM anorectic; antidiabetic; protein-protein interaction; diabetes;
XX	KM yeast 2-hybrid assay; metabolic disorder; obesity.
OS	Homo sapiens.
PN	WO200286122-A2.
PD	31-OCT-2002.
XX	14-MAR-2002; 2002MO-EP003768.
XX	14-MAR-2001; 2001US-0275734P.
PA	(HYBR-) HYBRIGENICS.
PI	Legrain P, Daviet L;
DR	WPI; 2003-103412/09.
XX	P-PSDB; AB070405.
PT	New complex between two interacting proteins in adipocyte cells, useful
PT	for identifying selected interacting domains that modulate protein
PT	interactions, or for preventing or treating metabolic disorders such as
PS	obesity or diabetes.
PS	Claim 7; Page 120-122; 382pp; English.
XX	The invention relates to a complex between two interacting proteins in
CC	adipocyte cells, given in the specification. The proteins are identified
CC	by selecting a bait protein from a known adipocyte marker and then
CC	performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC	members of an adipocyte cDNA library. The proteins are designated SID
CC	(RTM) (selected interacting domains) proteins. Also included are a
CC	polynucleotide encoding a polypeptide in the adipocyte cells, a
CC	recombinant host cell expressing at least one of the interacting
CC	polypeptides of the complex, selecting a modulating compound in adipocyte
CC	cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
CC	sequences given in the specification (including its fragment or variant),
CC	a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC	given in the specification (including its fragment or variant), a vector
CC	comprising the SID (RTM) polynucleotide, a recombinant host cell
CC	comprising the vector, a protein chip comprising the polypeptides and a
CC	record comprising all or part of the data, listed in the specification.
CC	The complex, polypeptides, polynucleotides and compounds are useful for
CC	preventing or treating metabolic disorders such as obesity or diabetes.
CC	The polynucleotides are useful as probes or primers. The complex is
CC	particularly useful for identifying selected interacting domains (SID
CC	(RTM)) for screening drugs that modulate the protein interaction, thus
CC	exhibiting the therapeutic effect. The present sequence encodes a SID
CC	(prey) protein of the invention
XX	Sequence 6007 BP; 1810 A; 1357 C; 1550 G; 1290 T; 0 U; 0 Other;
SO	Query Match 32.3%; Score 655.4; DB 8; Length 6007;
	Best Local Similarity 62.9%; Pred. No. 1.8e-148;
	Matches 1134; Conservative 0; Mismatches 645; Indels 24; Gaps 7;
QY	41 CCGGCCCCAGCCCTTGAATATGATGATCTTAAGCGAGCAGGAGAAATCGGCACTG 100
DB	1259 CCCTTCAGGCATGATGATGATGAGCAGCTCAGGCGAGCAACAAGAGGAATGAGGCAATTA 1318
QY	101 CGTAGGTGATGATGCTGAACAACAAGCTCATATATGATTAAGTAAGCAAAATCGGCC---A 156
DB	1319 AGGGATCTATATGCTGGAACACAAACCTCATATATGCAAACTACTAATAATAGGCCACAA 1378

QY	157	CAGTACTGAATGAGCCCTGGGGGAAAGCCTTTCTATCCAAAGAGATATGTCGACCGAC	216
Db	1379	CTAAAGGATTTAAACCCCTGAGAGAAAGGGAAATGGTGGAGAAATAATCCAAAGACGAA	1438
QY	217	ACCTTTTACAGTCACTTTAAAGAAAGATGTCAAAAAGCGTGTCTGGCACTGGATGAAGCC	276
Db	1439	AAACATGTATGCCAAATTAAGAGAGAGGTGGCCAGGAGCCCTGGCTCTGGATGAAAGCC	1498
QY	277	ATTCTCTCAATCAATC-----AGTTCCATGACAAATAGATTCAGATTCCTTGAAGC	327
Db	1499	GTGTCCCACTGCACACAGATTTACAGAGTTTCATATATAAAATTTGAGCCCTATGTTGAAGCA	1558
QY	328	CTGGAAACGCATCTGTGAAACGTCCTGAGGACAGCACCCCTCTATCTCTGCAGAGTTGAAG	387
Db	1559	CTGGAGATCTTTTCTCTCGCGCTCGTATGCGACACATGATCCCTGCTGAAGTAGACAG	1618
QY	388	ATCAAGAAACGATCACTGAAATAAGATGTGTCAGTAGACATGAAAGACTACAGCG	447
Db	1619	ATCAGAGAGTGCATCAGTACGAAATAAAGTCCACCGCTGAGCTTAGAAAACTCAGACCA	1678
QY	448	TTGATAGAACTCTTTAAACAGAGGGGAGAGGAATGATGCTATGATCTGGGGGGACTGAT	507
Db	1679	TCCTTTGAGCCTTGAAGCGCCGAGAGAGAGCTTATTTGACATCTCAGGGAGCAGAC	1738
QY	508	AAAGACATATCTGCCAAGCTGTTCAAGATTAAGCTTGACCAAAATGTTTTCATTTGGAG	567
Db	1739	AAGGATCTGGCTGCAAAAAGAAATCCAGATTAATTTGATCCAAATGTATTTCTTCGGAG	1798
QY	568	AAACATACACACACTGTGTGGAAGAGAGGAAGCCAACTACGTGATGTGATGACTACGA	627
Db	1799	GACATCAAAAGCTGGGCTGAAGAACGAGAAATCAAAATTTCTTGATGTCTTGAATTTGCA	1858
QY	628	GAAAAGTTCTGTGTGATCACTATGTCAATTGATTTACCATTAAGATACTCAAGATTTTC	687
Db	1859	GAGAAAGTTCTGTGTGATCACTATGTCAATTGATTTACCATTAAGATACTCAAGATTTTC	1918
QY	688	ATCCGGGACCTGGAAGATCTCGAATTGATCTTCAAGTAGTAAACACAGCAAGAACCA	747
Db	1919	GTCCATGACTTTGGAAAGCCGAGCATGTGATCTTTCATCATCAACAAACAGGTGGAAGCT	1978
QY	748	GCAGAGACATTAAGAGAAAGAAATGATGAGACTACAGAGGAGCGTGATATGTATTAATAC	807
Db	1979	GCTGAGACTATTAAGAGAGAGACAGATGTGTCTGATGAGAGCTGGAGTTTATTCGATTC	2038
QY	808	CTAGGTTCTGAACCTATTGCGGCATGTGGGAGCCTGATTAACCCATTTGTCMAAGAAGT	867
Db	2039	CTTGAGAGACATTTGATTTTGTCTGTGAGAAACTGAGAACTCGAAGTGAAGAGAGC	2098
QY	868	ATACATAGATTAAATTACAGCATTTGGATTCCTTAATTAAGCTTGGAAAGACCGGATTGA	927
Db	2099	ATTGATGAGATGAATTAATGTC--TTGGAGAACTTAACAAACATGAGAAAGAGAGCTAGA	2157
QY	928	CCAACTTGAG-----AGCAATGAGAGCTGCCGTCACTAGATGAGACTGACAGCGGATTT	981
Db	2158	AAAACCTTGAGAGATCTATGCAAGTGTCTGTGACATATCAGAGCACTCTTACAGCTATGTT	2217
QY	982	TTTCTGGGTAGATATTGCAAGTGTGTAAGTTTCCTTAATGTCTTCAATTGGAACAGCATC	1041
Db	2218	TGACTGGCTAGATTAACCTGTGATTAAACTCTGCACCATGTGCCCTGTGTGGCACTG--ACC	2276
QY	1042	TCGAATCGTCAACACAGCATTTGATGTGCTAGAGCAATTTAAGTCTGAGGCGTATACAC	1101
Db	2277	TCAAATCTGTTAAAGATCAAGTTAAATGAATGAAGAGCTTCAAGTGAAGATTACCAAC	2336
QY	1102	AGCAGATAGATGTGTACGACTGACTCATCAACGACAGCGTTTGTCTAAAGAAATTAACG	1161
Db	2337	AGCAATTTGAGATGAGAGAGCTTAATATCAACGAGGTGAACGTGATGTTAAAGAAATCTACTG	2396
QY	1162	NAGAGATGACAAACACACTGTTTATACACCAAAATAGANCTGATATTGATAT--GGCTA	1220
Db	2397	ATGAGACGAGACAGACATTTATAGAAACCACTGACAGAACTTCAACACCTCTGGAGAGA	2456

QY 1221 GCTGAGANNAGATCATCAACAGACGATAAAGGAGGCTCTATTAGACGCTG 1280
 DB 2457 ACCCTGGGTGAGAAATTTGCCACGACGACAACTAGAAAGGGCTCTTTGGCCCTTG 2516
 QY 1281 GTCAGTTCCACATGCCCCCTGTATAGCTCTGCAT--GCTGACACACACCGAGGCTTGC 1338
 DB 2517 GTCAGTTCCAGCATGCTTTAGAGAACTAATAGTTGGCTGCTACATACCAAGAGTGT 2576
 QY 1339 TAAATGACGACAAACCTGTGGAGAGACCTTAAAGCATTGAATTGAATCTGCCAAGC 1398
 DB 2577 TAGATGCTCAGAGACCAATAGTAGAGAGCCAAAAGTCAATTAAGTTCACCAAGC 2636
 QY 1399 ATCATGCTCTCAAAATGATGATATTAGCCATCACTGACAGTGAAGCCGTTATATAAG 1458
 DB 2637 ACCATGCTCTTAAAAATGATGATTTTGGCTCATCAAGCCACAGTGAACATCAACAAAG 2636
 QY 1459 CAGGAAATGATCTAATTTGAATCAAGTGCAGAGAGAAAGCAAGCAACCTTGACAAACAGC 1518
 DB 2697 CTGGCAATGAGCTTCTGAATCCAGTGTGAGATGATGCCAGCAGCTTAAGAGCCGTT 2756
 QY 1519 TAGAGGTTTAAATCAACGCTGCGCAAAATGTTTGGAAAAACAGACAAAGAACAGC 1578
 DB 2757 TGGAGGCCATGAAACCAATGCTGGAGAGTCAAGTTTACAGAAAAACAGAGAGAGGAGCAGC 2816
 QY 1579 AGCTGATGCTGCTGCTGGCCAGGCAAGGGTTCCATGGGAAATTGAGGATTTGCAAGC 1638
 DB 2817 AGCTTCAGTCACTCTGACAGAGGCCAGGCTTCCACAGTGAATGAAATTTCTCT 2876
 QY 1639 AGTGGCTGACTGACAGCAGGAGCTCATCTGTTGGCATCTTAAACCGCTGGAGGTTTACCGG 1698
 DB 2877 TGGAACTTACTAGATGAGAGAGCCAGCTTTCTGATCTTAACCCACAGAGAGACTTCCTG 2936
 QY 1699 AAACAGCCAGAGCAGCAGCTTAATGTCATATGAAAGTGTGCTGCTCTTGAAGCTAAAG 1758
 DB 2937 AAATGCTAGGGGAAACAGCTGTATACATATGAACTTATTTCCAGCTGAAGCCAAAGG 2996
 QY 1759 AAGAAACATTAAGAGTCTGATGACAAAGGCCAGACAGTCTTGCAAGATGCCAAAT 1818
 DB 2997 AAGGACTTATTAACAACACTGATGACAGGAGACATCATCTTAAAGCCGTGACGACT 3056
 QY 1819 CTG 1821
 DB 3057 CTG 3059
 RESULT 7
 AAS22908
 ID AAS22908 standard; DNA; 16122 BP.
 XX
 AC AAS22908;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE DNA encoding novel bone marrow polypeptide #2.
 XX
 KM Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
 KM haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
 KM wound healing; nutritional supplement; immune disorder;
 KM severe combined immunodeficiency; SCID; ds.
 XX
 OS Homo sapiens.
 XX
 PN MO200157187-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US003782.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 30-NOV-2000; 2000US-0250683P.
 XX

PA (HYSE-) HYSEQ INC.
 XX
 PI Ford JB, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
 PI Ren F, Drmanac RT;
 XX
 DR WPI; 2001-488875/53.
 DR P-PSDB; AAU14603.
 XX
 PT Nucleic acid encoding bone marrow polypeptides, useful in diagnostic and
 PT gene therapy.
 XX
 PS Claim 1; Page 146-151; 3922p; English.
 XX
 CC AAS22907-AAS23099 represent nucleic acids encoding novel bone marrow
 CC polypeptides. The nucleic acids and corresponding proteins may be used in
 CC the prevention, diagnosis and treatment of diseases associated with
 CC inappropriate bone marrow polypeptide expression. For example, to treat
 CC disorders associated with decreased expression by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of the
 CC polypeptides by expressing inactive proteins or to supplement the
 CC patient's own production of the polypeptide. Additionally, the nucleic
 CC acids may be used to produce the polypeptides, by inserting the nucleic
 CC acids into a host cell and culturing the cell to express the protein. The
 CC nucleic acid and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples, and therefore which patients
 CC may be in need of restorative therapy. The proteins may also be used as
 CC antigens in the production of antibodies against bone marrow proteins and
 CC in assays to identify modulators of their expression and activity. The
 CC anti-bone marrow protein antibodies and antagonists may also be used to
 CC down regulate expression and activity. The antibodies may also be used as
 CC diagnostic agents for detecting the presence of the protein in samples
 CC (e.g. by enzyme linked immunosorbent assay (ELISA)). The proteins may be
 CC used to regulate haematopoiesis activity, and consequently in the
 CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
 CC as wound healing; as a nutritional supplement; and in treatment of
 CC immune disorders such as severe combined immunodeficiency (SCID)
 XX
 SQ Sequence 16122 BP; 4917 A; 3572 C; 4169 G; 3464 T; 0 U; 0 Other;
 Query Match 32.3%; Score 655.4; DB 4; Length 16122;
 Best Local Similarity 62.9%; Pred. No. 2.6e-148;
 Matches 1134; Conservative 0; Mismatches 645; Indels 24; Gaps 7;
 QY 41 CCCGCCCGAGCCCTTGAATATGATCTCTAAGCAGCAGCAGGAAATATCGGCACTG 100
 DB 11377 CCCTCTCAGCAGTATCTAGACAGCTCGACGACAAACAGAGAAATGAGGCAATTA 11436
 QY 101 CGTAGTTGATAGCTGAACCAACGCTCATTAATGAATGAACAAACTGGCC---A 156
 DB 11437 AGGGAATCTATGCTGAACCAAACTCATATTTGACAAACTTAAGATAGAGCCACAA 11496
 QY 157 CAGTACTGAATGAGCCCTGGGGAAGGCTTTTCTATCCAGAGAAATATGTCGACCGGAC 216
 DB 11497 CTAAAGAAATTAACCTCTGAGAGAGGGAATGTGGAAGAAATATCCAGAAAGCAGAA 11556
 QY 217 ACCCTTACAGTCACATTAAGAAGATGTCAAAAAGCTGTGTGSCACTGATGAAGCC 276
 DB 11557 AACATGTATGCCCAATTAAGAGAGAGTGTGGCCAGAGGCCCTGTGATGAAGGCC 11616
 QY 277 ATTTCTCAATCAATC-----AGTTCAATACAAATAGATCAAGATCTTGAAGC 327
 DB 11617 GTGTCCAGTCCACAGATTAACAGAAATTTCTATTAATTAATGAGCCATATGTTGAAGCA 11676
 QY 328 CTGGAAGCATCTGTGAAGCTGTGAGGACAGCAGCTCTATCTTGTGAGAGAGTTGAAG 387
 DB 11677 CTGGAAGATCTTTCTCTCCCTGCGCTGATGACCACTGATCCCTGCTGAAGTAGACAG 11736
 QY 388 ATCAAGAACAGATCAGTGAATTAAGAAATGTGTCACTAGACATGAAAGAACTCAAGCCG 447
 DB 11737 ATCAGAGAGTGCATCAATTAAGATGAGCCACCGTGGAGCTAGAAAGAACTCAGCAGC 11796
 QY 448 TTGTATGAACCTTTAAACAGAGGGGAGAGAAATGATTTGCTAATCTGGGGGAGACTGAT 507

Db 11797 TCCTTGGAGGCTTTGAGAGCCGCTGGAGAGAGCTTAATGAGCATCTCAGGAGAGAC 11856
 Qy 508 AAGACATATCTGCGCAAGCTGTTCAAGATTAAGTCAACCAATGTTTTCTATTGGAG 567
 Db 11857 AAGGATCTGGCTCAAAAGAAATCCAGGATTAATGATCAAAATGGATCTTCTGGAG 11916
 Qy 568 AACATACACACACTGTTGAGAGAGAGAGAGCCAACTACTGATGTGAGACTAGCA 627
 Db 11917 GACATCAAAAGCTGGGGGTGAAGAACGAGAAATCAAAATTTCTGATGTCTTGAATTA 11976
 Qy 628 GAAAAGTTCTGGTGTGATCACATGTCAATGATGTTAATCAATTAAGTACTCAAGATT 687
 Db 11977 GAGAAATTTCTGGTATGACATGAGCAGCTCTCCAGCAACCTCAAGACACCCGAGAT 12036
 Qy 688 ATCCGGGACCTGGAAGTCCGGAATTGATCTTCACTAGTAAACAACAGCAAGAGCA 747
 Db 12037 GTCCATGACTTTGGAAGGCCAGGCAATGATCTTCAATCAACCAACAGTTGAAGCT 12096
 Qy 748 GCAGAGACCATTAAGAGAGAAATAGATGCACTACAGAGAGCTGATATATATTAAAC 807
 Db 12097 GCTGAGACTATTAAGAGAGAGAGAGAGATGCTGATGATGAAGAGCTGGAGTTATTC 12156
 Qy 808 CTAGGTTCTGAATCTATTCGCGCATGTGGGAGCTTGATTAACCCATTTGCAAGAAAG 867
 Db 12157 CTGGAGACGATTTGATTTTTCCTGGAGAACTGAGAAAGCTGAGAGAGAGAGAG 12216
 Qy 868 ATACATGAGTTAAATTCAGGAGATGCTCTTAATTAAGTCTGGAANAACCGGATTA 927
 Db 12217 ATTGATGAGATGAATTAATGC-TTGGAGAACTTAACCAAAACATGGAAGAGAGCT 12275
 Qy 928 CCAACTTGAAG-----AGCAATCAGCTGCTGCTGATACAGATGACCTGAGCGGAT 981
 Db 12276 AAAAATTGAGAGATGATAGCAAGCTGCTGAGATACAGGACACTTTCAGCTATGTT 12335
 Qy 982 TTTCTGGTATGATATTGACAGCTGTAACTTCTCAATGCTTCAATTTGAAACAGCAT 1041
 Db 12336 TGACTGGCTAGATTAACCTGTGATTAACCTGACACATGCCCTGTGGCAGTGC-ACC 12394
 Qy 1042 TCGAACTGTCAAGCAGAGATGATGATGCTAGAGCAATTAAGTCTGAGAGCTTACAC 1101
 Db 12395 TCAATCTGTTAAAGATCAGTTAAATGAAGAGGTTCAAGTAGAAGTTTACCAAC 12454
 Qy 1102 AGCAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1161
 Db 12455 AGCAATTTGAGATGAGAGAGCTTAATCAGCAGGCTGATGATTAAGAGAGCTACT 12514
 Qy 1162 MAGAGTGAACAACACTGTTCTAGACCAAAATGAGCTGATATGATAT-GGCTTA 1220
 Db 12515 ATGAGACGACAGAGACATTAATCGAGAACCTGACAGAACCTCAACCTCTGGGAG 12574
 Qy 1221 GCCTGAGNAGAGATCATCAACAGACAGATTAACCTGAGAGGCTGATTAAGCCNTG 1280
 Db 12575 ACTTGGGTGAGAAATTTGCCACGACAGCAAACTAGAGAGGCTGTGTGCCCTTG 12634
 Qy 1281 GTCAAGTCCAACTGCTGTATAGAGCTCTGCTGAT--GCTGACACACAGAGGCTTGC 1338
 Db 12635 GTCAAGTCCAACTGCTGTATAGAGAGCTTAATGAGTGTGCTGATCTACCGAAGAG 12694
 Qy 1339 TAAAGTGAAGAACTGTTGAGAGAGAGAGAGCCCTTAAGCCATTTGAATTTGAAC 1398
 Db 12695 TAGATGCTGAGAACCAATTAAGTGAAGACCAAAAGTCAATTTGAAGTTCGCAAG 12754
 Qy 1399 ATCATGTGCTCCAAATGATGATTAAGCCATAGTCCACAGTGAAGCCGTTAATTAAG 1458
 Db 12755 ACCATGTCTTAATAAATGATGTTTGTCTATCAAGCCACAGTGAAGAGTCAACAAAG 12814
 Qy 1459 CAGGAATGATCTAATGATCAAGTGAAGAGAGAGAGAGCAACCTTCAGAAACAG 1518
 Db 12815 CTGGCATATGAGCTTCTGATATCAGTGTCTGAGATGATGACAGCAGCTTAAGAG 12874
 Qy 1519 TAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAAAACAGAAACAAAGAGCAG 1578

Db 12875 TGGAAACATGACCAATCTGGAGTCACTGTTACAGAAAAACAGAGAGAGAGCAGC 12934
 Qy 1579 AGCTGATGATGCTTGGCCAGAGCCAAAGGTTCCATGCGCAATTTGAGATTTGCAGC 1638
 Db 12935 AGCTTATGATCACTTGCAGCAGGCTCCAGGCTTCCACATGTAATTTGAATTTCT 12994
 Qy 1639 AGTGGCTGACTGACAGAGAGCTCATCTGTGGCATCTAAACCGCTGGAGGTTTACCG 1698
 Db 12995 TGGAACTTAAGTAATGAGAGAGCAGCTTTCGATCTTAAGCCACAGAGAGACTTC 13054
 Qy 1699 AAACAGCCAGAGAGAGCTTAATGTCATATGGAAGTCTGTGCTTGAAGCTTAAG 1758
 Db 13055 AAATGCTTGGAGAACGCTTGAATACATATGGAATCTTAATCCAGCTGAAGCCAA 13114
 Qy 1759 AAGAAACATATAAGTCTGATGACAGAAAGCCAGAGATGCTGCAAGATGCCAAAT 1818
 Db 13115 AAGAGACTTAATTAATCAACTTGAACAGAGAGAGACTCAATGCTTGAAGCCG 13174
 Qy 1819 CTG 1821
 Db 13175 CTG 13177
 RESULT 8
 AAS23002
 ID AAS23002 standard; DNA; 16341 BP.
 XX
 AC AAS23002;
 DT 24-OCT-2001 (first entry)
 XX
 DE DNA encoding novel bone marrow polypeptide #96.
 XX
 KW Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
 KW haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
 KW wound healing; nutritional supplement; immune disorder;
 KW severe combined immunodeficiency; SCID; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157187-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US003782.
 XX
 PR 03-FEB-2000; 2000US-0046914.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 30-NOV-2000; 2000US-0250683P.
 XX
 PA (HYBR-) HYBRQ INC.
 XX
 PI Ford JE, Boyle BJ, Tang YF, Liu C, Auundi V, Zhou P, Xue AJ;
 PI Ren F, Drianaac RT;
 PI Ren F, Drianaac RT;
 DR MPI: 2001-488875/53.
 XX
 PT P-PSDB; AAUI4697.
 PT
 PT Nucleic acids encoding bone marrow polypeptides, useful in diagnostic and
 gene therapy.
 PS Claim 1; Page 253-257; 392pp; English.
 XX
 AAAS22907-AAS23009 represent nucleic acids encoding novel bone marrow
 CC polypeptides. The nucleic acids and corresponding proteins may be used in
 CC the prevention, diagnosis and treatment of diseases associated with
 CC inappropriate bone marrow polypeptide expression. For example, to treat
 CC disorders associated with decreased expression by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of the
 CC polypeptides by expressing inactive proteins or to supplement the
 CC patient's own production of the polypeptide. Additionally, the nucleic
 CC acids may be used to produce the polypeptides, by inserting the nucleic

CC acids into a host cell and culturing the cell to express the protein. The
CC nuclear acid and its complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and therefore which patients
CC may be in need of restorative therapy. The proteins may also be used as
CC antigens in the production of antibodies against bone marrow proteins and
CC in assays to identify modulators of their expression and activity. The
CC anti-bone marrow protein antibodies and antagonists may also be used to
CC down regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the protein in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). The proteins may be
CC used to regulate haematopoiesis activity, and consequently in the
CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
CC such as wound healing; as a nutritional supplement; and in treatment of
CC immune disorders such as severe combined immunodeficiency (SCID)
XX
SQ Sequence 16341 BP; 4965 A; 3637 C; 4241 G; 3498 T; 0 U; 0 Other;
Query Match 32.3%; Score 655.4; DB 4; Length 16341;
Best Local Similarity 62.9%; Pred. No. 2.6e-148;
Matches 1134; Conservative 0; Mismatches 645; Indels 24; Gaps 7;
QY 41 CCCGCCCCAGCCCTTGTAATATGATATCTTAAGCAGCAGCAGAGAAATCGGCACTG 100
DB 11596 CCGCTCCAGCCATTGATCATGAGCAGCTCAGGAGCAACAGAGAAATAGGCAATTA 11655
QY 101 CGTGAGTTGATAGCTGAACACAGCCTCATATAGTAAGTGAACAAACCTGGCC---A 156
DB 11656 AGGAAATCTATGTGGAACACAAACCTCATATTTGCAAACTAATAATATGAGCCCAAA 11715
QY 157 CAGTACTGAATGAGCCCTGGGGAAGGCTTTCTATCCAGAGAAAGTATGTGGACCGAG 216
DB 11716 CTAAGAGAAATTAACCTCGAGAAAGGGGAAATGGTGAAGAAATATCCAAAGACAA 11775
QY 217 ACCCTTACAGTCACTTAATAAGAGATGTCAAAAACGCTGCTGTGCACTGGAATGAAGC 276
DB 11776 AACATGATAGCCCAATTAAGAGAGAGGTGGCCAGCAGCCCTGTGCTTGGATGAAGCC 11835
QY 277 ATTTCTCAATCACTC-----AGTTCCATGACAAAGATAGATCAGATCCTTGAAGC 327
DB 11836 GTGTCCCAAGTCCACAGATTAACAGAGTTTCATGATTAATAATTGAGCCATATGTTGAGACA 11895
QY 328 CTGGAACCCATCTGTGGAACCTTCAAGCAGCAGCCCTTATCTCTGCAAGAGTTGAAGAG 387
DB 11896 CTGGAAGATCTTCTCTCTGCTGTGCTGTGCACTGATCCCTGCTGAAGTGAACAAAG 11955
QY 388 ATCAAGAGAACAGATCAGTGAATAATTAAGATGTGTCACTGATGATGAAGAACTACAGCCG 447
DB 11956 ATCAAGAGTGCATCACTGAGACATTAAGAGTCCACCGTGAAGCTAGAAAACCTGACGCA 12015
QY 448 TTGTATGAACTCTTAACAGAGGGAGAGAAATGATTGCTAGATCTGGGGGGACTGAT 507
DB 12016 TCCCTTGGAGCCTTGAAGCGCGCTGAGAGAGAGCTTATTGACATCTCAAGGAGACAGAC 12075
QY 508 AAACACATATCTGCCAAAGCTGTTCAGAGATAAGCTTGAACAAATGCTTTTCAATTTGGAG 567
DB 12076 AAGGATCTGGCTGCAAAAGAAATCCAGATTAATTTGATCAAAATGATATTTCTTGGAG 12135
QY 568 AACATACACACACTGTGTGAAGAGAGAGCAAACTACTGATGTGATGAGAGCTACGA 627
DB 12136 GACATCAAAAGCTCGGGCTGAAGAGAGAAATCAAAATTTCTTGATGTCTTGAATTAACA 12195
QY 628 GAAAAGTTCTGGTGTGATCAGATGTGATTAAGTATTAACATTAAGATTAAGATTTTC 687
DB 12196 GAGAAAGTTCTGGTGTGATCAGATGTGATTAAGTATTAAGTATTAAGATTTTC 12255
QY 688 ATCCGGGACCTGGAAGATCTGGAATTGAATCTTCAAGTAAAGCAACACAGAGAAACA 747
DB 12256 GTCCATGACTTGAAGAGCCCGAGCTTATCTTCAATCAATCAACAGAGTTGAAGCT 12315
QY 748 GCAGAGACCATTAAGAGAGAAATAGATGACTACAGAGAGAGCTGATATATTAATTAAC 807
DB 12316 GCTGAGACTATTAAGAGAGAGACAGATGTGTCTGCAATGAAGAGCTGAGATTAATTCGATC 12375

QY 808 CTAGCTTCTGAACCTATTGCGGCAATGCGGAGGAGCTGATTAACCCATTTGTCAAGAGAT 867
DB 12376 CTTGAGACAGATTTGATTTTGTCTGTGGAGAAACAGAGAGCTGAAGTGAAGAGAGC 12435
QY 868 ATACATAGATTAATTAATGAGGATGAGATTTCTTAATAATTAAGCTTGAAGAACCCGATTTGA 927
DB 12436 ATTATGATGATTAATTAATGCT--TTGGAGAACTTAACAAACATGAAAGAGAGGCTAGA 12494
QY 928 CCAACTTGAAG----AGCAATGACGCTCCGTCAGTACAGATGAGTCAAGCGCGATTT 981
DB 12495 AAACTGAGGATGCTATGCAAGCTGTGAGCAATACAGACACTCTTCAGCTATGCTT 12554
QY 982 TTTTGGGTGATATTTGCAAGTGTGATGCTTCAATGTCTCAATTTGGAACAGATC 1041
DB 12555 TGACTGTGATTAACCTGTGATTAACCTTGACACATGCCCCCTGTGGCACTG--ACC 12613
QY 1042 TCGAACTGTCAACAGACAGATTAATGATGTGCTAGAGCAATTAATTAAGTGAAGCTATCAAC 1101
DB 12614 TCAATATCTTTAAGATCACTTAATTAATGAAGAGGCTTCAAGTGAAGATTTTACCAAC 12673
QY 1102 AGCAGATAGAGTGTGACACTGACTCATCAAGCAGACGCTTTTGTAAAGAACTAACAG 1161
DB 12674 AGCAAAATTGAGATGAGAGAGCTTAATATCACCAGGGTGAATGTATTAAGAAAGCTACTG 12733
QY 1162 MAGAGATGACAAACACACTGTTTGAACCCAAATATGAGTATATTAAT--GGCTA 1220
DB 12734 ATGAGAGCGAGAGAGATTAATTAAGAAACCACTGAAGAACTCAACACCTCTGGAGAGA 12793
QY 1221 GCCTGAGAGAGAGATCATCAACAGACAGCTAACTGAGAGGCTCTATTAAGCCATG 1280
DB 12794 ACCGTGGTGAAGAAATTTGCCACAGACAGCAACAACTGAAGAGGCTCTGTGGCCCTTG 12853
QY 1281 GTCAGTTCCAACATGCCCTGTATGAGCTCTGCAAT--GCTGACACACACCGAGGCTTGC 1338
DB 12854 GTCACTTCCAGATGCTTATGAGAGAACTAATGATGTGGCTGATCATATCCGAAAGATTGT 12913
QY 1339 TAAGTGAACGAAACCTGTTGAGAGAGACCTTAAGCCATTTGAATTTGAATCTTCCCAAGC 1358
DB 12914 TAGATGTCTGAGAGCCAAATTAAGTGAAGACCCAAAGTCAATTTGAATGAGTCTCCAAAGC 12973
QY 1399 ATCATGTCTCCAAATGATATTAAGTATTAAGCCATGATCCACAGTGAAGAGCGTTAATAAG 1458
DB 12974 ACCATGTCTTAATAATGATGATTTGTGCTCATCAAGCCACAGTGAAGAAACGTCAACAAAG 13033
QY 1459 CAGGAAATGATCTTAATTAATCAAGTCAAGAGAGAGAAAGCAAGCACTTCAGAAACAAGC 1518
DB 13034 CTGGCAATGAGCTTCTTGAATTCAGTCTGAGATGATGACAGCAGCTTAAGAGAGCGTT 13093
QY 1519 TAGAGTTTAAATCAACGCTGCAAAATGTTTGGAAAAACAGAAACAAAGAGACAGC 1578
DB 13094 TGAAGGCACTGAACCAATGCTGAGAGTCAAGTGTATTCAGAAAAACAGAGAGGAGAGCAGC 13153
QY 1579 AGCTGAGATGAGTCTTGGCGCAGGSCCAAGGCTTCCATGGCGAAATTGAGAGATTGACAGC 1638
DB 13154 AGCTTCACTCAACTCTGAGAGAGGCCAGGCTTCCAGAGAAATTAAGATTTCTCT 13213
QY 1639 AGTGGCTGACTGACACAGGAGCTGATCTGTTGGCATCTAAACCGCTGGAGGTTTACCGG 1698
DB 13214 TGAAGATTAATGAATGAGAGAGCAGCTTCTGATCTTAAGCCACAGAGAGATTTCTGTG 13273
QY 1699 AAACAGCCAAAGAGCAGCTTAATGTCAATATGAGAGTGTGTGCTGCTTTGAAGCTAAAG 1758
DB 13274 AAATCTGTAGGAAACACCTTGAATCAATATGAACTCTATTTCCACAGCTGAAGCCAGG 13333
QY 1759 AAGAAATATTAAGATCTGATGAGAAAGGCGCAGAGATGCTTGAAGAGATGCCAAAT 1818
DB 13334 AAGGAGCTTATTAACACTACTGACAGAGGAGACATCAATGCTTCTTAAGCGGTGACGACT 13393
QY 1819 CTG 1821
DB 13394 CTG 13396

RESULT 9
ID ABK34676 standard; cDNA; 2181 BP.
XX
XX AC ABK34676;
XX
XX 08-MAY-2002 (first entry)
XX
XX Human cDNA for novel secreted protein, SEQ ID 445.
XX
XX Human; ss; gene; secreted protein; immune deficiency; viral infection;
XX bacterial infection; fungal infection; autoimmune disorder; burn;
XX rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
XX diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
XX Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
XX coagulation disorder; haemophilia; inflammatory disorder; ulcer;
XX tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
XX lymphoid cell deficiency.
XX
XX Homo sapiens.
XX
XX MO200177290-A2.
XX
XX 18-OCT-2001.
XX
XX 29-MAR-2001; 2001MO-US010295.
XX
XX 06-APR-2000; 2000US-0194941P.
XX
XX (GBMY) GENETICS INST INC.
XX
XX Wong GG, Clark HF, Fecthel K, Agostino MJ, Howes SH, Resnick RJ;
XX PI Gulukota K, Graham JF;
XX
XX WPI; 2002-179323/23.
XX
XX Six hundred and twenty five polynucleotides derived from a variety of
XX human tissue sources which encode secreted proteins, useful for treating
XX immune deficiencies and disorders such as autoimmune disorders.
XX
XX
XX Claim 1; Page 234; 339pp; English.
XX
XX The invention relates to 625 polynucleotides which have been derived from
XX a variety of human tissue sources and sequences that hybridise to them. Also
XX proteins, their complements and sequences that hybridise to them. Also
XX included are a vector comprising the polynucleotide, a host cell
XX transformed with the vector, the proteins encoded by the polynucleotides,
XX antibodies that bind to the proteins and identification of modulators of
XX the proteins or the expression of the polynucleotide. The polynucleotides
XX can be used as probes for the identification and isolation of full length
XX cDNA and genomic DNA. The polynucleotides and proteins can also be used
XX as nutritional supplements. The protein is useful in the treatment of
XX various immune deficiencies and disorders such as viral infections,
XX bacterial infections, fungal infections, autoimmune disorders (e.g.
XX rheumatoid arthritis, multiple sclerosis, autoimmune thyroiditis and
XX diabetes) and allergic reactions and conditions (e.g. asthma). They are
XX also useful for treating neurodegenerative diseases (e.g. Alzheimer's
XX disease, Parkinson's disease), liver fibrosis, coagulation disorders
XX (e.g. haemophilia), inflammatory disorders (e.g. Crohn's disease) and
XX tumours. They are also useful for tissue regeneration, for wound healing
XX and in the treatment of burns, incisions and ulcers. The proteins are
XX also useful for regulating haematopoiesis, for treating myeloid or
XX lymphoid cell deficiencies. The present sequence is one of the 625 cDNA
XX sequences encoding a secreted protein
XX
XX Sequence 2181 BP; 679 A; 442 C; 545 G; 515 T; 0 U; 0 Other;
XX
XX Query Match 24.7%; Score 501; DB 6; Length 2181;
XX Best Local Similarity 62.0%; Pred. No. 3,7e-111;
XX Matches 895; Conservative 0; Mismatches 529; Indels 19; Gaps 6;
XX
XX 296 TCATGACAAAGATGATCAGATCCTTGAGAGCCTGGAAACGATCGTGGAAAGCTTGAGGC 355

Db 1 TTCATGATAAATTTGAGCCTATGTTGAGACACTGGAGATCTTTCTCGCTCCGTA 60
Qy 356 AGCCACCTCTATCTCTGAGAGGTTGAGAAAGTCAAGAAACGATGAGTAAATTAAG 415
Db 61 TGCCACCTAGTATCCTGCTGAAAGTACAAAGATCAGAAAGTCAATGATGACATTAAG 120
Qy 416 ATGTGTCACTAGACATGAGAAAGCTACAGCCGTTGTATGAAACTTTAAACAGAGGAG 475
Db 121 GTGCCACCGTGGAGTAGAAGAAAATCGACCATCTTTGAGGCGTTGAAGCCGCTGAG 180
Qy 476 AGGAATGATTTGCTAGATCTGGGGGACCTGATTAAGACATATCTGCCAAGCTGTTCAG 535
Db 181 AGAGGTTATTTGAGACATCTCAGGGAGCAGACAAAGATCTGGTGAAGAAATCAAG 240
Qy 536 ATAAAGTTAACCAATGTTTTCATTTGAGAGAACATACACACATCTGTGAGAGAGAGG 595
Db 241 ATTAATTTGATCAATATGATTTCTTTGGAGAGACATCAAAAGTCCGGCTGAAGAACAG 300
Qy 596 AAGCCAACTACTGATGATGATGAGCTAGACAGAAAAGTTCTGTGTGATCAATGTCAAT 655
Db 301 AATCAAAATTTCTGATGATCTTGATATGACAGAAAGTTCTGTGTGATGACATGGCAGCTC 360
Qy 656 TGATAGTTACCATTAAGATTAATCAAGATTTTATCCGGGACCTGGAAGATCTGGAATTG 715
Db 361 TCTGACACACATCAAAAGACACCCAGGATTTGTCCATATCTTGGAAACCCAGGCAATGG 420
Qy 716 ATCCCTCAGTAGTAAACAAACAGCAAGAGAGAGACAGACATTAAGGGAAGAAATGATG 775
Db 421 ATCTTTCATCATCAACAAACAGGTTGAAAGCTGTGAGACTTTTAAGAAAGACAGATG 480
Qy 776 GACTACAGAGAGAGCTGATATATGATTAATTAACCTAGTTCTGAACATCTATGGCAGTGTG 835
Db 481 GTCTGATGAAGAGCTGAGATTATTCGATCCGATCTTGAGACAGATTTGATTTTGGCTGTG 540
Qy 836 GGGAGCTGATTAACCCATTGTCAAGAAAGATTAATCATAGTTAAATTAAGCAGTGGAT 895
Db 541 GAGAACTAGAGAGCCTGAAATGAATTAATGCTTGGAGAACTTAAACGAAACATG--- 597
Qy 896 TCTCTAAATTAAGCTTGGAAAGCCGAGATGACAACTTGAGAGCAATGAGTGGCCTG 955
Db 598 ---AAGATAGGCTAGCAAAACCTTG-----AGATGCTATGAGGCTGTGTGAG 646
Qy 956 AGTACAGATGAGCTGAGAGCGGATTTTCTGGTATGATTAATTCAGAGTGAATGCTGCT 1015
Db 647 TATCAGAGAGCTCTTCAAGCTATGTTGACTGCGAAAGATTAACATGTTAACTCTTC 706
Qy 1016 TCAATGCTTCAATTTGAAACAGATCTGAAACTGTCAAGCAGCAGATTTGATGCTTAA 1075
Db 707 ACCAAGCCCCCTGTGGCTCTG-ACCTCGATCTGTTAAAGATCAAGTTAAATGAATGA 765
Qy 1076 GCATTTAAGTCTGAGGCGCTATCAACAGAGATGAGTGGTGCAGTCACTATCAACGA 1135
Db 766 GAGGTTCAAGTGAAGATGTACACGACGAAATGATGAGAGAACTTAAATCAACGAGG 825
Qy 1136 GAGGCTTTTGTAAAGAAAGTAACAGNAGAGAGTGAACAAACACTGTTCTAGACCCAAA 1195
Db 826 TGAAGTATGATTAAGAAAGCTAAGATGATGAGACGACAGACATTAATACGAAACCACT 885
Qy 1196 AATGAGCTGATTAATGATATGCT-AGCTGAGNAGAGAAATCATCAAGACAGATTA 1254
Db 886 GACAGAACTCAAAACACTCTGAGAGAACTGGGTGAGAAATTTGCCACCGCAGACAGAA 945
Qy 1255 ACTGAGGGGTGCTTATTTAGCCNTGGGTGAGTTTCAACATGCGCTGATAGACTCTGCA 1314
Db 946 ACTGAGGGGCTCTGTGGCCCTTGTGATGTTCCAGCATGCTTTAGAGAACTAATGAG 1005
Qy 1315 T--GCTGACACACAGAGGCTTGTGTAAGTGAAGAAACCTGTTGAGAGAGACCTTAA 1372
Db 1006 TTGGCTGATCTATACGAGAGAGTGTGATGATCTCTAGAGAACCAATTAAGTGAAGACCAA 1065
Qy 1373 AGCCATTGAATTTGAACCTTGCAGAGCATGTGCTCAAAATGATGATTAAGCCATCA 1432

PR	6-JUL-2000	2000US-02209633
PR	26-JUL-2000	2000US-02209649
PR	14-AUG-2000	2000US-02245188
PR	14-AUG-2000	2000US-02245198
PR	14-AUG-2000	2000US-02252133
PR	14-AUG-2000	2000US-02252149
PR	14-AUG-2000	2000US-02252667
PR	14-AUG-2000	2000US-02252677
PR	14-AUG-2000	2000US-02252688
PR	14-AUG-2000	2000US-02252707
PR	14-AUG-2000	2000US-02254477
PR	14-AUG-2000	2000US-02254577
PR	14-AUG-2000	2000US-02257588
PR	18-AUG-2000	2000US-02267598
PR	18-AUG-2000	2000US-02267797
PR	22-AUG-2000	2000US-02266818
PR	22-AUG-2000	2000US-02266868
PR	22-AUG-2000	2000US-02271828
PR	30-AUG-2000	2000US-02289248
PR	01-SEP-2000	2000US-02292878
PR	01-SEP-2000	2000US-02293438
PR	01-SEP-2000	2000US-02293448
PR	05-SEP-2000	2000US-02293508
PR	05-SEP-2000	2000US-02293518
PR	06-SEP-2000	2000US-02304378
PR	06-SEP-2000	2000US-02304388
PR	08-SEP-2000	2000US-02312428
PR	08-SEP-2000	2000US-02312438
PR	08-SEP-2000	2000US-02312448
PR	08-SEP-2000	2000US-02314138
PR	08-SEP-2000	2000US-02314148
PR	08-SEP-2000	2000US-02320808
PR	08-SEP-2000	2000US-02320818
PR	12-SEP-2000	2000US-02319688
PR	14-SEP-2000	2000US-02323378
PR	14-SEP-2000	2000US-02323388
PR	14-SEP-2000	2000US-02323398
PR	14-SEP-2000	2000US-02324008
PR	14-SEP-2000	2000US-02324018
PR	14-SEP-2000	2000US-02330638
PR	14-SEP-2000	2000US-02330648
PR	14-SEP-2000	2000US-02330658
PR	21-SEP-2000	2000US-02342238
PR	21-SEP-2000	2000US-02342748
PR	25-SEP-2000	2000US-02349978
PR	25-SEP-2000	2000US-02349988
PR	26-SEP-2000	2000US-02349588
PR	26-SEP-2000	2000US-02354848
PR	27-SEP-2000	2000US-02358348
PR	27-SEP-2000	2000US-02358358
PR	29-SEP-2000	2000US-02363278
PR	29-SEP-2000	2000US-02363678
PR	29-SEP-2000	2000US-02363688
PR	29-SEP-2000	2000US-02363698
PR	29-SEP-2000	2000US-02363708
PR	13-OCT-2000	2000US-02393938
PR	13-OCT-2000	2000US-02393978
PR	20-OCT-2000	2000US-02409608
PR	20-OCT-2000	2000US-02412218
PR	20-OCT-2000	2000US-02417858
PR	20-OCT-2000	2000US-02417868
PR	20-OCT-2000	2000US-02417878
PR	20-OCT-2000	2000US-02418088
PR	20-OCT-2000	2000US-02418098
PR	20-OCT-2000	2000US-02418268
PR	01-NOV-2000	2000US-02446178
PR	08-NOV-2000	2000US-02464748

Db 900 GAGACCAATAAGTGAGAGACCCAAAAGTCATTGAAGTGGAGCTCGCAAAAGCACATGTCCT 959
Qy 1409 CCAAAATGATATATTAGCCCATCAGTCCACAGTGGAGCCGTTAATAAGCAGAAATGA 1468
Db 960 AAAAATGATATTTTGGCTCATCAAGCCAGTGAACCAAGTCAGCAAAAGCTGGCAATGA 1019
Qy 1469 TCTAATTAATCAATGAGAGAGAAAGAACCAACCTTGAGAACAGTAGAGCTTT 1528
Db 1020 GCTTCTTAATCACTGCTGTGAGATGATGCCAGCAGCTTTAAGAGCCCTTTGAAAGCAT 1079
Qy 1529 AAATCAAGCTGCGCAAAATGTTTGGAAAAACAGAACAAAGAGACAGACGTGATGG 1588
Db 1080 GAACCAATGCTGGAGTCAAGTGTACAGAAAAACAGAGAGAGGAGCAGCAGCTTCACTC 1139
Qy 1589 TGCCTTGGCCAGGCGCAAAAGGTTCCATGGCGAAATTGAGATTGGCAGCAGTGGCTAC 1648
Db 1140 AACTCTGAGAGAGCCGAGGCGCTTCACAGTGAATTGAAGTTTCTCTTGGAACTTAC 1199
Qy 1649 TGACACGAGGAGTCACTGTGGCATTAACCGGTGGAGGTTTACCGGAAACAGCCAA 1708
Db 1200 TAGAATGAGAGAGCCAGCTTTCTGCATCTAAGCCACAGAGAGACTTCTGAAACTGTAG 1259
Qy 1709 GAGCAGCTTAATGTCATATGAGAGTGTGCTGCTTTGAGCTAAAGAAACATA 1768
Db 1260 GGAACAGCTTGATACATATGGAATCTATTTCCAGCTGAAGCAAGAGAAAGACTTA 1319
Qy 1769 TAAGAGTGTGATGCAGAAAGCCAGACAGATGCTTGCAAGATGCCAAATCTG 1821
Db 1320 TAATCACTACTTGACAGAGGCGAGACTCATGCTTAAAGCCGTGACGACTCTG 1372

RESULT 11

AA529791 ID AA529791 standard; cDNA; 5453 BP.

AC AA529791;

DT 21-NOV-2001 (first entry)

XX Human cytoskeletal element-related polypeptide encoding cDNA #22.

KW Cytoskeletal element-related protein; human; mouse; rabbit; goat; horse;
cat; dog; chicken; sheep; immunosuppressive; antirheumatic; vasotropic;
antirheumatic; antiproliferative; cytoskeletal; cardiac; neuroprotective;
cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
ophtalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;
cerebrovascular disorder; nervous system disorder; bacterial infection;
fungal infection; viral infection; ocular disorder; endocrine disorder;
gastrointestinal disorder; renal disorder; respiratory disorder;
wound healing; skin aging; organ transplantation; food preservative;
tissue regeneration; anti-fertility; food additive.

XX Homo sapiens.

XX WO200151568-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001331.

XX PF 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.

20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251888P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251900P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX MPI: 2001-476182/51.
DR P-PSDB; AUI08529.
XX
PT Novel isolated human cytoskeletal element-related polypeptide useful for
PT diagnosis/treatment of neoplastic disorders, disorders associated with
PT neural transmission, chromosomal abnormalities, autoimmune disorders.
XX
PS Claim 1; SEQ ID NO 32; 505pp; English.

CC hyperproliferative disorders such as neoplasms of the breast or liver,
CC cardiovascular disorders such as cardiac arrest, cerebrovascular
CC disorders such as cerebral ischemia, nervous system disorders such as
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC ocular disorders such as corneal infection, endocrine disorders such as
CC premature labour and infertility, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma. The polypeptides can also be used
CC to aid wound healing, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, to regenerate tissues and in chemotaxis.
CC The polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at

Query Match 24.5%; Score 497.4; DB 4; Length 5453;
Best Local Similarity 63.7%; Pred. No. 4e-110;
Matches 835; Conservative 0; Mismatches 465; Indels 11; Gaps 5;

QY 520 GCCAAGCTGTCAGGATTAAGCTTGACCAATGGTTTCAATTGGGAGAACATACACACA 579
DB 1 GCAAAAGAAATCCAGATTAATGGATCAATGGTATTTCTGGGAGACATCAAGCT 60
QY 580 CTGGTGAAGAGAGGGAGCCAACTACTGATGTGATGAGCTACAGAAAATTCTCG 639
DB 61 CGGGCTGAAGAGAGGAATCAAAATTTCTGATGTCTTGAATTGACAGAAAGTTCTGG 120
QY 640 TGTGATCAGATGTCATTGATTAAGTACCAATTAAGATCTCAAGATTTCACGGGACCTG 699
DB 121 TATGACATGGCAGCTCTCTCTGACCAATCAAGACCCAGGATTTGTCATGACTTG 180
QY 700 GAAGATCCCTGGAATGATGCTTCTCAGATGAATTAACACAGACAGACAGACACATA 759
DB 181 GAAAGCCAGGCTTGATGATCTTCATCATCAACCAAGATTTGACCTGAGACTTAT 240
QY 760 AGGAGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 819
DB 241 AAGGAAGACACAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 820 CTCATTGCGGATGCGGAGGAGCTGATTAACCATTTGCAAGAAAGTATACATGATGAT 879
DB 301 TTGATTTTGGCTGCTGAGAGAACTGAGAACTGAGAACTGAGAACTGAGAACTGAGAA 360
QY 880 AATTACGATGGGATTTCTTAATTAAGCTTGAAGAACCGGATTTGACCACTTGA-- 937
DB 361 AATTATGC-TTGGAGAACTTAACCAACATGAGAAAGAGGCTTGAAGAACTTGAAGA 419
QY 938 ---AGCAATGAGCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 993
DB 420 TCGTATGCAAGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
QY 994 TATTGAGGTGTAAGTGTCTTCAATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 1053
DB 480 TAACACTGTGATTAATCTGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 538
QY 1054 AGCAGAGATGATGCTGAGCAATTAAGTCTGAGGCTTATCAACGACGATGATGAT 1113
DB 539 AAGATCAAGTAAATGAATGAAGAGCTTCAAGTGAAGTTTACCAACGCAATTTGAGA 598
QY 1114 GGTAGACCTGATCTATCAGCAGAGCTTTTCTAAAGAAATTAACAGNAGAGATGACA 1173
DB 599 TGGAGAGCTTAATACACAGGCTGAAGTATTTAAAGAACTGATGATGAGACGAGACA 658
QY 1174 AACACACTGTTTAAGCCCAAAATGAGATGATATTTGATAT-GGCTAGCTTGAAGNAGA 1232
DB 659 GAGACATTATACAGAAACACAGACAGAACTCAAAACCTCTGGGAGAACTGGGAGAGA 718
QY 1233 GAATATCAACAGACAGATTAACGAGGCTCTATTAACCTGCTGCTGCTGCTGCTGCTG 1292
DB 719 AATATGCCACGAGACAGCAACATGAAGGAGGCTCTGTTGGCCCTTGTGCTGCTGCTGCT 778
QY 1293 ATGCCCTGATAGCTCTGCTGAT--GCTGACACACACGAGGCTTGTAACTGAGACAGA 1350

Db 779 ATGCTTAGAGGAAGTAATGATGGCTGACTCATACCGAAGAGTTGTTAGATGCTCAGA 838
Qy 1351 AACCTGTTGAGAGAGACCTTAAGCCATTGAAATTGAACCTTGCCAGACATATGCTCC 1410
Db 839 GACCAATAAGGGAGACCCAAAGATCTATGTAAGTTGAGCTGCAAGACCATGCTCTAA 898
Qy 1411 AAAATGATGTTATTTAGCCCATGATGCCACAGTGGAAAGCCGTTAATTAAGACAGAAATGATC 1470
Db 899 AAAATGATGTTATTTAGCCCATGATGCCACAGTGGAAAGCCGTTAATTAAGACAGAAATGATG 958
Qy 1471 TAAATGATCATGATGACAGAGAAAGCAAGCAACTTCAGAAACAGCTAGAGGTTTAA 1530
Db 959 TTCTTGAATCCAGTGTGAGATGATGCCAGACCTTAAGAGCCGTTGGAAGCCATGA 1018
Qy 1531 ATCAACCTGAGCAAAATGTTTGAAGAAACAGAAACAAAGAGCAGAGCTGATGCTG 1590
Db 1019 ACCAATCTGGGAGTCAAGTGTTAAGAAACAGAGAGAGGAGCAGACGCTTCAAGTCAA 1078
Qy 1591 CCTTGCCGCGAGCCAAAGGGTTTCATGCGCAATTGAGATTTGACAGAGTGGCTGACTG 1650
Db 1079 CTCTGACAGAGGCCCGAGGGCTTCCACAGTGAATTTGAATTTCTTGGAACTTACTA 1138
Qy 1651 ACAGAGAGCGTCACTGTGGCATCTAAACCGCTGGAGGTTTACCGGAACAGCCAAAG 1710
Db 1139 GAATGAGAGGCACTTCTGCACTTAAGCCACAGAGAGACTTCTGAACTGCTAGGG 1198
Qy 1711 AGCAGCTTAATGTCATATGGAAGTGTGCTGCTTGAAGCTTAAGAGAAACATATA 1770
Db 1199 AACAGCTGTATACATATGTGAAGTCTATTTCCAGCTGAAGCAAGAGAAAGACTTATA 1258
Qy 1771 AGACTGTGATGCAAGAAAGGCCAGAGAGTGTGCAAGATGCCCAAAATCTG 1821
Db 1259 ATCAACTACTTGAACAAGGCCAGACTCATGCTTTTAAGCCGTGAGACTCTG 1309

RESULT 12

AAL22375 standard; cDNA; 547 BP.

AAL22375;

07-DEC-2001 (first entry)

Human breast cancer expressed polynucleotide 14832.

Human; breast cancer; cell marker; cytostatic; ss.

Homo sapiens.

WO200151628-A2.

19-JUL-2001.

10-JAN-2001; 2001WO-US000798.

14-JAN-2000; 2000US-0176077P.

14-MAR-2000; 2000US-0189167P.

24-MAR-2000; 2000US-0192099P.

29-MAR-2000; 2000US-0193480P.

15-MAY-2000; 2000US-0205230P.

09-JUN-2000; 2000US-0211315P.

25-JUL-2000; 2000US-0220534P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Lillie J, Xu Y, Wang Y, Steinmann K;

WPI; 2001-451856/48.

New peptide useful as a marker for the diagnosis of breast cancer.

Claim 1; Page 2676; 3695bp; English.

CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX

SQ Sequence 547 BP; 186 A; 103 C; 140 G; 118 T; 0 U; 0 Other;

Query Match 21.6%; Score 438.8; DB 4; Length 547;

Best Local Similarity 94.1%; Pred. No. 2.8e-96;

Matches 508; Conservative 0; Mismatches 26; Indels 6; Gaps 5;

Qy 962 GATGACTGACGCGGATTTTCTGGGTAGATATTCAGAGTGTAGATTGCTTCAATG 1021

Db 10 GATGACTGACGCGGATTTTCTGGGTAGATATTCAGAGTGTAGATTGCTTCAATG 69

Qy 1022 TCTCAATTGGAACAGCATCTCGAACTGTCAAGCAGCAGATTGATGCTAGACAAAT 1081

Db 70 TCTCAATTGGAACAG-ATCTCGAACTGTCAAGCAGCAGATTGAAAGCTTAAAGCAATT 128

Qy 1082 TAACTGAGGCGCTATCAAGCAGATAG-AGTGTAGACTGATCATCAGCAGAGCGC 1140

Db 129 TAACTGAGGCGCTATCAAGCAGATAGAAATGGAAGACTGATCATCAAGCAGAG-GC 187

Qy 1141 TTTTGCTAAAGAAAGTAAACAGNAGAGATGACAAACACACTGTTCTAGACCCAAATG 1200

Db 188 TTTTGCTAAAGAAAGTAAACAGNAGAGATGACAAACACACTGTTCTAGACCCAAATG 247

Qy 1201 AMCTGATATTTGAT-GGTAGCTGGAAGAGATATCATCAAGCAGATTAATCTGG 1259

Db 248 AACTGAATTTGATATGAGATAGCCTGGAAGAGAAATCATCAAGCAGATTAATCTGG 307

Qy 1260 AGGTGCTCTATTAGCCGTCAGTTCACAACTGCGTGAAGACTCTG--CATGC 1317

Db 308 AGGTGCTCTATTAGCCGTCAGTTCACAACTGCGTGAAGACTCTGCGATGCG 367

Qy 1318 TGACACACACGAGGCTGCTTAAGTGAAGCAAACTGTTGAGAGAGACCTTAAGCCA 1377

Db 368 TGACACACACGAGGCTGCTTAAGTGAAGCAAACTGTTGAGAGAGACCTTAAGCCA 427

Qy 1378 TTGAATTTGAACCTTCCAAAGCATATGCTCCAAATGATGATTAAGCCCATGATCCA 1437

Db 428 TTGAATTTGAACCTTCCAAAGCATATGCTCCAAATGATGATTAAGCCCATGATCCA 487

Qy 1438 CAGTGAAGCGCTTAATTAAGCAGAAATGATCTTAATTTGAATCAAGTGCAGGAAGAAG 1487

Db 488 CAGTGAAGCGCTTAATTAAGCAGAAATGATCTTAATTTGAATCAAGTGCAGGAAGAAG 547

RESULT 13

AAL13506 standard; cDNA; 531 BP.

AAL13506;

07-DEC-2001 (first entry)

Human breast cancer expressed polynucleotide 5963.

Human; breast cancer; cell marker; cytostatic; ss.

Homo sapiens.

WO200151628-A2.

19-JUL-2001.

10-JAN-2001; 2001WO-US000798.


```

Db 208 GGGAGAGAAATGATTGCTAGATCTGGGGGAGCTGATTAACAATATGTGCAAGCTGC
Qy 531 TCAGATTAAGCTTGAACCAATAGTTTTCATTGGGAGAACATACACACTGTGGAGA 590
Db 268 TCAGATTAAGCTTGAACCAATAGTTTTCATTGGGAGAACATACACACTGTGGAGA 327
Qy 591 GAGGGAAGCCAACTACTGATGTGATGAGCTAGCAGAAAGTTTGTGTGATTCAT 650
Db 328 GAGGGAAGCCAACTACTGATGTGATGAGCTAGCAGAAAGTTTGTGTGATTCAT 387
Qy 651 GTCATTGATAGTT 663
Db 388 GTCATTGATAGTT 400

RESULT 15
AAK51484
ID AAK51484 standard; cDNA; 4327 BP.
XX
AC AAK51484;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide seq ID NO 29.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00653561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Aeundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhtman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
DR P-PSDB; AAM78351.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 1; Page 570-575; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, hematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

```

```

CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 4327 BP; 1365 A; 921 C; 1046 G; 995 T; 0 U; 0 Other;
Qy Query Match 17.3%; Score 352.2; DB 4; Length 4327;
Best Local Similarity .94.9%; Pred. No. 6.5e-75;
Matches 374; Conservative 0; Mismatches 19; Indels 1; Gaps 1;
Qy 1494 GAGCAAGCAACCTTCAGAACAGCTAGAGCTTTTAATCAAGCTGGCAAAATGTTTG 1553
Db 1 GACAGAGCAACCTTCAGAACAGCTAGAGCTTTTAATCAAGCTGGCAAAATGTTTG 60
Qy 1554 GAAAAACAGAAACAAAGAGAGCAGACAGCTGATGCTGGCGCCAGGCAAAAGGCTTC 1613
Db 61 GAAAAACAGAAACAAAGAGAGCAGACAGCTGATGCTGGCGCCAGGCAAAAGGCTTC 119
Qy 1614 CATGGCGAAATTGAGATTTGACAGCACTGCTGACACGAGAGCTCATCTGTTGGCA 1673
Db 120 CATGGCGAAATTGAGATTTGACAGCACTGCTGACACGAGAGCTCATCTGTTGGCA 179
Qy 1674 TCTAACCGCTGGAGGTTTACCGGAAACAGCCAGAGAGAGCTTAATGTCATATGAA 1733
Db 180 TCTAACCGCTGGAGGTTTACCGGAAACAGCCAGAGAGAGCTTAATGTCATATGAA 239
Qy 1734 GTCTGTGCTGCTTGAAGCTTAAGAGAAACATATAAGTCTGATGCAAGAAAGCCAG 1793
Db 240 GTCTGTGCTGCTTGAAGCTTAAGAGAAACATATAAGTCTGATGCAAGAAAGCCAG 299
Qy 1794 CAGATGCTTGCAGATGCCCAAAATCTGCAGAGCAATATTGACCAAGACATTAATTAAC 1853
Db 300 CAGATGCTTGCAGATGCCCAAAATCTGCAGAGCAATATTGACCAAGACATTAATTAAC 359
Qy 1854 TTGAAAAAATGGAATCGGTGAAACCAAC 1887
Db 360 TTGAAAAAATGGAATCGGTGAAACCAAC 393

```

Search completed: January 12, 2005, 08:50:38
 Job time : 932 secs

inis Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2005, 11:10:16 : Search time 1086 Seconds
(without alignments)
10745.764 Million cell updates/sec

Title: US-10-089-887-46

Perfect score: 2031
Sequence: 1 ctgcacatgcgcagcagaaca.....tngagagaccacacgcgc 2031

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2024	99.7	2031	15	US-10-089-887-46
2	1656.8	81.6	9274	9	US-09-885-535-3
3	1653.6	81.4	16684	18	US-10-723-860-7690
4	1637.6	80.6	3307	16	US-10-108-2604-1453
5	807	39.7	3987	14	US-10-198-846-12468
6	652.2	32.1	18529	14	US-10-198-846-12599
7	501	24.7	2181	9	US-09-823-245A-445
8	497.8	24.5	2113	15	US-10-158-034-52
9	497.4	24.5	5453	15	US-10-158-034-32
10	430.8	21.2	599	14	US-10-198-846-4738
11	266.8	13.1	4495	15	US-10-037-270-352
12	266.8	13.1	4495	15	US-10-117-722-352

13	262.8	12.9	344	11	US-09-864-408A-111	Sequence 111, App
14	235.6	11.6	1721	14	US-10-043-487-126	Sequence 126, App
15	177.8	8.8	3684	15	US-10-094-749-272	Sequence 272, App
16	134.6	6.6	563	15	US-10-029-386-4198	Sequence 4198, App
17	133.6	6.6	171	15	US-10-029-386-17898	Sequence 17898, App
18	93.8	4.6	1024	13	US-10-202-193-9	Sequence 9, App1
19	88.6	4.4	449	15	US-10-158-034-51	Sequence 51, App1
20	88.6	4.4	4156	15	US-10-158-034-31	Sequence 31, App1
21	65.2	3.2	3720	10	US-09-822-846-180	Sequence 180, App
22	64.4	3.2	845	14	US-10-043-487-125	Sequence 125, App
23	63.4	3.1	364	14	US-10-198-846-9302	Sequence 9302, App
24	63.2	3.1	2117	15	US-10-104-047-16	Sequence 16, App1
25	61.2	3.0	361	14	US-10-198-846-8468	Sequence 8468, App
26	61	3.0	1595	15	US-10-369-493-28935	Sequence 28935, App
27	61	3.0	1595	15	US-10-369-493-28937	Sequence 28937, App
28	60	3.0	60	10	US-09-808-975-12022	Sequence 12022, App
29	58.8	2.9	3510	10	US-09-845-416-12	Sequence 12, App1
30	58.8	2.9	4476	10	US-09-845-416-31	Sequence 31, App1
31	58.6	2.9	663	18	US-10-425-115-86406	Sequence 86406, App
32	58.2	2.9	923	16	US-10-343-710-16	Sequence 16, App1
33	57.6	2.8	450	15	US-10-312-495-4	Sequence 4, App1
34	57.6	2.8	2048	10	US-09-814-353-20207	Sequence 20207, App
35	56.8	2.8	361	14	US-10-198-846-8619	Sequence 8619, App
36	56.2	2.8	14568	15	US-10-311-455-203	Sequence 203, App
37	55.2	2.7	139	18	US-10-425-115-80204	Sequence 80204, App
38	55.2	2.7	716	9	US-09-772-134B-81	Sequence 81, App1
39	54	2.7	2357	15	US-10-106-698-2037	Sequence 2037, App
40	53.4	2.6	1352	13	US-10-080-960-7	Sequence 7, App1
41	53	2.6	1947	15	US-10-171-581-135	Sequence 135, App
42	52.8	2.6	674	16	US-10-168-948B-5	Sequence 5, App1
43	52.8	2.6	1629	18	US-10-425-115-97514	Sequence 97514, App
44	52.8	2.6	2519	13	US-10-139-262-9	Sequence 9, App1
45	52.8	2.6	2519	14	US-10-253-969-9	Sequence 9, App1

ALIGNMENTS

RESULT 1

US-10-089-887-46

Sequence 46, Application US/10089887

Publication No. US20030219740A1

GENERAL INFORMATION:

APPLICANT: Bayer Corporation et al.

TITLE OF INVENTION: DNA Sequences Isolated from Human Colonic Epithelial Cells

FILE REFERENCE: 1657/1020

CURRENT APPLICATION NUMBER: US/10/089, 887

CURRENT FILING DATE: 2000-08-08

PRIOR APPLICATION NUMBER: US 60/147, 933

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 61

SOFTWARE: PatentIn version 3.1

SEQ ID NO 46

LENGTH: 2031

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (1162)..(1162)

OTHER INFORMATION: Unknown

FEATURE:

NAME/KEY: misc feature

LOCATION: (1202)..(1202)

OTHER INFORMATION: Unknown

FEATURE:

NAME/KEY: misc feature

LOCATION: (1277)..(1277)

OTHER INFORMATION: Unknown

FEATURE:

/ NAME/KEY: misc feature
/ LOCATION: (1885) .. (1885)
/ OTHER INFORMATION: Unknown
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1889) .. (1889)
/ OTHER INFORMATION: Unknown
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (2013) .. (2013)
/ OTHER INFORMATION: Unknown
US-10-089-887-46

Query Match 99.7%; Score 2024; DB 15; Length 2031;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2031; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGCGATGCTGACAGAAACCAATCAATCTCTCAGCTCCGCCGCCAGCCCTTGAAAT 60
DB 1 TTGCGATGCTGACAGAAACCAATCAATCTCTCAGCTCCGCCGCCAGCCCTTGAAAT 60
QY 61 TGAATCTTAAGGACAGAGAGAGAAACATTCGGCAATCGCTGAGTGTATAGCTGAACA 120
DB 61 TGAATCTTAAGGACAGAGAGAGAAACATTCGGCAATCGCTGAGTGTATAGCTGAACA 120
QY 121 CAAGCTCATATAGATTAAGTGAACAAACTGGCCACAGTACTGAATGAGCCCTGGGGAA 180
DB 121 CAAGCTCATATAGATTAAGTGAACAAACTGGCCACAGTACTGAATGAGCCCTGGGGAA 180
QY 181 GGGCTTTCTATCCAGAGAAAGTATGTGGCAGCCGACACCTTTTACAGTCAATTAAGAA 240
DB 181 GGGCTTTCTATCCAGAGAAAGTATGTGGCAGCCGACACCTTTTACAGTCAATTAAGAA 240
QY 241 GATGTCAAAAAGGCTGTGGCACTGGAGTGAAGCCATTTCTCAATCACTCACTTCCT 300
DB 241 GATGTCAAAAAGGCTGTGGCACTGGAGTGAAGCCATTTCTCAATCACTCACTTCCT 300
QY 301 GACAAATAGATCAGATCCTTTGAGAGCTTGAAGCAGTCTGGAAAGCTGTGAGGACCA 360
DB 301 GACAAATAGATCAGATCCTTTGAGAGCTTGAAGCAGTCTGGAAAGCTGTGAGGACCA 360
QY 361 CCTCTATCTCTGACAGAGTTGAGAAAGATCAAGAAACAGATCACTGAAATTAAGTGT 420
DB 361 CCTCTATCTCTGACAGAGTTGAGAAAGATCAAGAAACAGATCACTGAAATTAAGTGT 420
QY 421 TCAGTGAACAATGGAAGGCTAGACCGGTGTATGAACCTTTTAAACAGAGGGAGAGAA 480
DB 421 TCAGTGAACAATGGAAGGCTAGACCGGTGTATGAACCTTTTAAACAGAGGGAGAGAA 480
QY 481 ATGATTGCTAGATCTGGGGGGAAGCTGATTAAGACATATCTGCCAAAGCTGTTAGAGTA 540
DB 481 ATGATTGCTAGATCTGGGGGGAAGCTGATTAAGACATATCTGCCAAAGCTGTTAGAGTA 540
QY 541 CTGGACCAAAATGGTTTTTCAATTTGGGAGAACATACACACTGGTGAAGAGAGAGGCC 600
DB 541 CTGGACCAAAATGGTTTTTCAATTTGGGAGAACATACACACTGGTGAAGAGAGAGGCC 600
QY 601 AAACCTAGTGAATGATGAGAGCTAGAGAAAGTTTGTGTGTATCAACATGTATTTGATA 660
DB 601 AAACCTAGTGAATGATGAGAGCTAGAGAAAGTTTGTGTGTATCAACATGTATTTGATA 660
QY 661 GTTACATTAAGATTAAGTCAAGATTTCAATCCGGGACCTGGAAAGATCTGGAAATTTGATCCT 720
DB 661 GTTACATTAAGATTAAGTCAAGATTTCAATCCGGGACCTGGAAAGATCTGGAAATTTGATCCT 720
QY 721 TCAGTATGAAGCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 TCAGTATGAAGCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 CAGGAGAGAGCTGATTAAGTATTAAGTCTAGAGTTCCTGAACTCATTTGGGAGAGAG 840
DB 781 CAGGAGAGAGCTGATTAAGTATTAAGTCTAGAGTTCCTGAACTCATTTGGGAGAGAG 840

QY 841 CCTGATAAACCATTTGCTCAAGAGAGATATAGATTAATTCAGCGATGGGATTTCTCT 900
DB 841 CCTGATAAACCATTTGCTCAAGAGAGATATAGATTAATTCAGCGATGGGATTTCTCT 900
QY 901 AAATTAAGCTTGGAAAGACCGGATTTGACCAACTTGAAGCAATGACGCTGCCGTAGTAC 960
DB 901 AAATTAAGCTTGGAAAGACCGGATTTGACCAACTTGAAGCAATGACGCTGCCGTAGTAC 960
QY 961 AGATGAGCTGACAGGCGGTATTTTTCTGGGTAGATATTTGAGAGGTGAAGTGGCTTCAAT 1020
DB 961 AGATGAGCTGACAGGCGGTATTTTTCTGGGTAGATATTTGAGAGGTGAAGTGGCTTCAAT 1020
QY 1021 GTCTCCAAATTTGAAACAGATCTCGAAACTGTCAAGAGAGAGATTTGCTAGAGCAAT 1080
DB 1021 GTCTCCAAATTTGAAACAGATCTCGAAACTGTCAAGAGAGAGATTTGCTAGAGCAAT 1080
QY 1081 TTAAGTCTGAGGCTTATCAACAGAGATAGAGTGTGACGACTGATCAACGACAGCCG 1140
DB 1081 TTAAGTCTGAGGCTTATCAACAGAGATAGAGTGTGACGACTGATCAACGACAGCCG 1140
QY 1141 TTTTGTAAAGAAAGTAAAGNAGAGAGTGAACAACACTGTCTTGAAGCCCAAAATGAG 1200
DB 1141 TTTTGTAAAGAAAGTAAAGNAGAGAGTGAACAACACTGTCTTGAAGCCCAAAATGAG 1200
QY 1201 ANCTGATATTGATATAGCTAGCTGGAGNAGAGAAATCATCAACAGACATTAATCTGA 1260
DB 1201 ANCTGATATTGATATAGCTAGCTGGAGNAGAGAAATCATCAACAGACATTAATCTGA 1260
QY 1261 GGGTGTCTATTAAGCCNTGGGTCAAGTCCCAATGCTCTGTATGAGCTCTGTGATCTGA 1320
DB 1261 GGGTGTCTATTAAGCCNTGGGTCAAGTCCCAATGCTCTGTATGAGCTCTGTGATCTGA 1320
QY 1321 CACACACGAGGGCTTGTCTAAGTGAAGAGAAACCTGTGAGAGAGACCTTAAGGCCATTG 1380
DB 1321 CACACACGAGGGCTTGTCTAAGTGAAGAGAAACCTGTGAGAGAGACCTTAAGGCCATTG 1380
QY 1381 AAATTAAGCTTCCAAAGCATCATGTGCTCCAAATGATATTTAGCCCATCACTGACAG 1440
DB 1381 AAATTAAGCTTCCAAAGCATCATGTGCTCCAAATGATATTTAGCCCATCACTGACAG 1440
QY 1441 TGAAGCCGTTAATTAAGAGAGAAATGATCTAATTTGAATCAAGTCAAGAGAGAGAGAA 1500
DB 1441 TGAAGCCGTTAATTAAGAGAGAAATGATCTAATTTGAATCAAGTCAAGAGAGAGAGAA 1500
QY 1501 GCAACCTTGAAGAAAGCTAGAGGTTTAAACACAGCTGGGAGAAATGTTTTGAGAAAAA 1560
DB 1501 GCAACCTTGAAGAAAGCTAGAGGTTTAAACACAGCTGGGAGAAATGTTTTGAGAAAAA 1560
QY 1561 CAGAACAAAGAGAGAGAGCTGATGTGCTTGGCCAGGCAAGGGTTCCATGAGCG 1620
DB 1561 CAGAACAAAGAGAGAGAGCTGATGTGCTTGGCCAGGCAAGGGTTCCATGAGCG 1620
QY 1621 AAATTAAGATTTGACAGAGTGTGCTGATCTGACAGAGAGGTCATCTGTTGGAATCAAC 1680
DB 1621 AAATTAAGATTTGACAGAGTGTGCTGATCTGACAGAGAGGTCATCTGTTGGAATCAAC 1680
QY 1681 CGCTGGAGAGTTTACCGGAAAGAGCCAGCAAGCAAGCTTAATGCAATGAGAGCTGTG 1740
DB 1681 CGCTGGAGAGTTTACCGGAAAGAGCCAGCAAGCAAGCTTAATGCAATGAGAGCTGTG 1740
QY 1741 CTGCTTTGAAAGCTAAAGAGAAATATTAAGTCTGATGAGAGAGAGAGAGAGAGATG 1800
DB 1741 CTGCTTTGAAAGCTAAAGAGAAATATTAAGTCTGATGAGAGAGAGAGAGAGAGATG 1800
QY 1801 TTGCAAGATGCCCAAAATCTGACAGAGACAAATATTTGACCAAGACATTAATCTTGA 1860
DB 1801 TTGCAAGATGCCCAAAATCTGACAGAGACAAATATTTGACCAAGACATTAATCTTGA 1860
QY 1861 AAATTAAGCTTGGAAAGACCGGATTTGACCAACTTGAAGCAATGACGCTGCCGTAGTAC 1920
DB 1861 AAATTAAGCTTGGAAAGACCGGATTTGACCAACTTGAAGCAATGACGCTGCCGTAGTAC 1920
QY 1921 AATTTAGACCTGCGCGGCGGCTGAGACCTTATAGTGAAGTGTATTTAGAGATGAA 1980


```

QY      1673  ATCTAAACCGGTGGAGGTTTACCGGAAACGCAAGAGAGCGTTAATGTCATATGGA 1732
Db      5289  ATCTAAACCGGTGGAGGTTTACCGGAAACGCAAGAGAGCGTTAATGTCATATGGA 5348
QY      1733  AGCTGTGCTGCTCTTTGAAGCTAAAGAAAGAAATATTAAGTGTGAATGCAGAAAGCCA 1792
Db      5349  AGTGTGTGCTGCTCTTTGAAGCTAAAGAAAGAAATATTAAGTGTGAATGCAGAAAGCCA 5408
QY      1793  GCAGATGCTTGCAAGATGCCCCAAATGTCGAGAGACAAATTTTGAACCAAGACATTAATTA 1852
Db      5409  GCAATGCTTGCAAGATGCCCCAAATGTCGAGAGACAAATTTTGAACCAAGACATTAATTA 5468
QY      1853  CTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAC 1887
Db      5469  CTTGAAAGAAAAATGGGAAATCGTGGAAACCAAC 5503

RESULT 3
US-10-723-860-7690
; Sequence 7690, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natacha
; APPLICANT: Ginsburg, Wendy M.
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882,0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7690
; LENGTH: 16684
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (857)..(877)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1454)..(1488)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1960)..(2049)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3121)..(3143)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4680)..(4703)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6993)..(7011)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10498)..(10533)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13384)..(13422)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13686)..(13756)

```

```

; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15350)..(15368)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-7630

Query Match      81.4%; Score 1653.6; DB 18; Length 16684;
Beet Local Similarity 95.1%; Pred. No. 0;
Matches 1803; Conservative 0; Mismatches 74; Indels 18; Gaps 9;

QY      8      GCTGACGAGAAACCAATCAATCATCTCTCAGC-TCCCGCCCGACGCCCTTGAAATGATATC 66
Db      10732  GCTGACGAGAAACCAATCAATCATCTCTCAGCCTTCCCGCCCGACGCCCTTGAAATGAAAC 10791

QY      67      TCTAAGGCAGCAGCAGAGAAACATCGGCACCTGCGTGAATTGATAGCTGAACACAAGCC 126
Db      10792  TCTAAGGCAGCAGCAGAGAAACATCGGCACCTGCGTGAATTGATAGCTGAACACAAGCC 10851

QY      127     TCATATGATATGATGATGACAAACTGG---CCACAGTACTGAATGAGCCCTGGGGAGG 182
Db      10852  TCATATGATATGATGATGACAAACTGGGGCCACAGTTACTGAATTGAGCCCTGGGGAGG 10911

QY      183     CTTTTCATCTCAGAGAAAGATGTGGCAGCCGACCCCTTACAGCACATTAAGAAAGA 242
Db      10912  CTTTTCATCTCAGAGAAAGATGTGGCAGCCGACCCCTTACAGCAAATTAAGAAAGA 10971

QY      243     TGTCAAAAAGCGTGTGTGCACTGGATAGGCCATTTCTCAATCAACTCAAGTTCATGA 302
Db      10972  TGTCAAAAAGCGTGTGTGCACTGGATAGGCCATTTCTCAATCAACTCAAGTTCATGA 11031

QY      303     CAAGATGATCAGATCTCTTGAGAGCTGGAAAGCATCGTGAACGTCTGAGGCAGCCACC 362
Db      11032  CAAGATGATCAGATCTCTTGAGAGCTGGAAAGCATCGTGAAGGTTTGAAGCAGCCACC 11091

QY      363     CTCATATCTCGCAGAGGTTGAAGATCAAGAAACAGATCAATGAAATTAAGATGTGC 422
Db      11092  CTCATATCTCGCAGAGGTTGAAGATCAAGAAACAGATCAATGAAATTAAGATGTGC 11151

QY      423     AGTAGACATGAAAAAGCTACAGCCGTTGTATGAACTCTTAAACAGAGGGAGAGAAAT 482
Db      11152  AGTAGACATGAAAAAGCTACAGCCGTTGTATGAACTCTTAAACAGAGGGAGAGAAAT 11211

QY      483     GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAGCTTTCAAGATPACT 542
Db      11212  GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAGCTTTCAAGATPACT 11271

QY      543     TGACCAAATGCTTTTCAATTTGGGAGAACATACACACTGCTGTGAAAGAGAGGAAGCAA 602
Db      11272  TGACCAAATGCTTTTCAATTTGGGAGAACATACACACTGCTGTGAAAGAGAGGAAGCAA 11331

QY      603     ACTACTGATGTGATGAGACTAGCAGAAAGTTTGGTGTGATCACATGCTCAATTATAGT 662
Db      11332  ACTACTGATGTGATGAGACTAGCAGAAAGTTTGGTGTGATCACATGCTCAATTATAGT 11391

QY      663     TACCATTAAAGATACTCAAGATTTCATCCGGGAGCTGGAAGATCTGTGAATTGATCTTTC 722
Db      11392  TACCATTAAAGATACTCAAGATTTCATCCGGGAGCTGGAAGATCTGTGAATTGATCTTTC 11451

QY      723     AGTAGAGTAAACCAACACAGAGAGAGCAGACAGACATTAAGGGAAGAAATAGATGACTCA 782
Db      11452  AGTAGAGTAAACCAACACAGAGAGAGCAGACAGACATTAAGGGAAGAAATAGATGACTCA 11511

QY      783     GAGAGAGCTGATATATGTTATTTAACTTAGTCTGAACTCATTTGGGAGCTGTGGGAGCC 842
Db      11512  GAGAGAGCTGATATATGTTATTTAACTTAGTCTGAACTCATTTGGGAGCTGTGGGAGCC 11571

QY      843     TGAATAAACCAATGTGCAGAGAGGTATACATGATGTTAAATTCAGCGATGGGATTTCTTAA 902
Db      11572  TGAATAAACCAATGTGCAGAGAGGTATACATGATGTTAAATTCAGC-ATGGGATTTCTTAA 11630

QY      903     ATTAAGCTTGGAAAGACCGGATTTGACCAATTGAG-----AGCAATGCACTGCCGTCA 956
Db      11630  ATTAAGCTTGGAAAGACCGGATTTGACCAATTGAG-----AGCAATGCACTGCCGTCA 11690

```

Db 11631 ATAAAGCTTGGAAAGACCGGATGTGCAAACTTGAAGAGCAATGACAGGCTGCCCTTCACT 11690
QY 957 GTACAGATGAGCTGACAGCGGATTTTCTGGCTTGAATTTGCGAGGTGGTAAGTTCCGTT 1016
Db 11691 ACCAGGATGAGCTGACAGCGGATTTTGAAGTGGATGATTTGAGTGGTGTAAATTAAGCTT 11750
QY 1017 CAATGCTCTCAATTTGGAACAGCATCTCGAAACTGTCAAGACGACAGATTGATGTGTAGAG 1076
Db 11751 CAATGCTCTCAATTTGGAACAGCATCTCGAAACTGTCAAGACGACAGATTGATGTGTAGAG 11809
QY 1077 CAATTTAAGTGTAGGCTTATCAACAGCAGATAG-AGTGTATGCACTGACTATCATCA 1135
Db 11810 CAATTTAAGTGTAGGCTTATCAACAGCAGATAGAAATGGAAGAAGCTGAATCATCAACGA 11869
QY 1136 GACGCTTTTGTCTAAAGAAAGTACAGNAGAGTGAACAACACACTGTTTGTAGACCCAA 1195
Db 11870 GA-CTTTTGTCTAAAGAAAGTACAGNAGAGTGAACAACACTGTTTCAAGACCCCAAT 11928
QY 1196 AATGGANCTGATTTATGATAT-GGCTAGGCTGGAAGNAGAAATCATCAACAGACAGCATTA 1254
Db 11929 AATGGANCTGATTTATGATATGATAGCTTGGAGAGAAATCATCAACAGACAGCATTA 11988
QY 1255 ACTGAGGCTGCTCTATTAGCCNTGGTCACTTCCACATGCCCCCTGTATGAGCTCTG-- 1312
Db 11989 ACTGAGGCTGCTCTATTAGCCNTGGTCACTTCCACATGCCCCCTGTATGAGCTCTG-- 12048
QY 1313 CATGTGACACACACCGAGGCTTGTATGAGACGAAACCTGTTGGAGAGACCTTAA 1372
Db 12049 ATGCTGACACACACCGAGGCTTGTATGAGACGAAACCTGTTGGAGAGACCTTAA 12108
QY 1373 AGCCATTGAAATTTGAATCTTGCACAGCATCATGTGCTCCAAAATGATGATTAAGCCATCA 1432
Db 12109 AGCCATTGAAATTTGAATCTTGCACAGCATCATGTGCTCCAAAATGATGATTAAGCCATCA 12168
QY 1433 GTCCACAGTGAAGCCGTTAAATTAAGCAGAAATGATCTAAATTTGATCAATGTCAGAGA 1492
Db 12169 GTCCACAGTGAAGCCGTTAAATTAAGCAGAAATGATCTAAATTTGATCAATGTCAGAGA 12228
QY 1493 AGAAGCAAGCACTTTCAGAACAACTAGAGTTTAAATCAACGCTGGCAAAATGTTTT 1552
Db 12229 AGAAGCAAGCACTTTCAGAACAACTAGAGTTTAAATCAACGCTGGCAAAATGTTTT 12288
QY 1553 GGAAGAAACAGAACAAAGAGAGCAGACGTGATGCTTGGCCAGGCAAGGCTT 1612
Db 12289 GGAAGAAACAGAACAAAGAGAGCAGACGTGATGCTTGGCCAGGCAAGGCTT 12348
QY 1613 CCATGCGCAAAATTTGAGATTTGACAGAGTGGCTGATGACACGAGAGCTCATCTGTTGGC 1672
Db 12349 CCATGCGCAAAATTTGAGATTTGACAGAGTGGCTGATGACACGAGAGCTCATCTGTTGGC 12408
QY 1673 ATCTPAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGCAGCTTAAATGTCATATGGA 1732
Db 12409 ATCTPAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGCAGCTTAAATGTCATATGGA 12468
QY 1733 AGTCTGCTGCTCTTGAAGCTTAAAGAAACATATAGAGCTCTGATGCGAAGAGGCA 1792
Db 12469 AGTCTGCTGCTCTTGAAGCTTAAAGAAACATATAGAGCTCTGATGCGAAGAGGCA 12528
QY 1793 GCAGATGCTTGCAGATGCTCCAAAATTTGACAGAGCAAAATTTGACCAAGCATTAATTA 1852
Db 12529 GCAGATGCTTGCAGATGCTCCAAAATTTGACAGAGCAAAATTTGACCAAGCATTAATTA 12588
QY 1853 CTTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1887
Db 12589 CTTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 12623

RESULT 4
US-10-108-260A-1453
; Sequence 1453, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE

; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108, 260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1453
; LENGTH: 3307
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1453
Query Match 80.6%; Score 1637.6; DB 16; Length 3307;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 1778; Conservative 0; Mismatches 58; Indels 18; Gaps 9;
QY 8 GCTGACAGAAACACATCAATCATCTCTACG-TCGCCGCCAGCCCTTGATATGATAC 66
Db 1457 GCTGACAGAAACACATCAATCATCTCTACGCTTCCGCCCCAGCCCTTGATATGAAAC 1516
QY 67 TCTAAGCAGCAGCAGAGAAACATCGGCAATCGCTGATGATTGATGCTGAACACAGCC 126
Db 1517 TCTAAGCAGCAGCAGAGAAACATCGGCAATCGCTGATGATTGATGCTGAACACAGCC 1576
QY 127 TCAATATGATATGATGAACAAACTGG---CCAAGTACTGATGAGCCCTGGGGAAG 182
Db 1577 TCAATATGATATGATGAACAAACTGGCCCAAGTACTGATGAGCCCTGGGGAAG 1636
QY 183 CTTTCTATCCAAAGAGATATGTGGACGCGACACCCCTTACAGTCAATTAAGAAGA 242
Db 1637 CTTTCTATCCAAAGAGATATGTGGACGCGACACCCCTTACAGTCAATTAAGAAGA 1696
QY 243 TGTCAAAAGCGTCTGTGCACTGATGAAGCCATTTCTCATCAATCTCAGTTCCATGA 302
Db 1697 TGTCAAAAGCGTCTGTGCACTGATGAAGCCATTTCTCATCAATCTCAGTTCCATGA 1756
QY 303 CAAGATGATCAGATCCTTGAGAGCCGGAACGATGCTGAGAGGCGACACC 362
Db 1757 CAAGATGATCAGATCCTTGAGAGCCGGAACGATGCTGAGAGGCGACACC 1816
QY 363 CTCTATCTCTGACAGGTTGAGATCATCAAGAACAGATCAGTGAATTAAGATGTCTC 422
Db 1817 CTCTATCTCTGACAGGTTGAGAGATCATCAAGAACAGATCAGTGAATTAAGATGTCTC 1876
QY 423 AGTACATGAGAAAGCTACAGCCGTTGATGAACCTTTAAACAGAGGAGAGAAAT 482
Db 1877 AGTACATGAGAAAGCTACAGCCGTTGATGAACCTTTAAACAGAGGAGAGAAAT 1936
QY 483 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGGCCAAAGCTTTCAGATTAAGCT 542
Db 1937 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGGCCAAAGCTTTCAGATTAAGCT 1996
QY 543 TGACCAATAGGTTTCTATTGGGAGACATACACACTGGTGAAGAGAGGAGCA 602
Db 1997 TGACCAATAGGTTTCTATTGGGAGACATACACACTGGTGAAGAGAGGAGCA 2056
QY 603 ACTACGATGATGATGAGCTAGCAGAAAGTTCTGCTGATGATCACTGCTATGATAGT 662
Db 2057 ACTACGATGATGATGAGCTAGCAGAAAGTTCTGCTGATGATCACTGCTATGATAGT 2116
QY 663 TACCATTAAGATATCTCAAGATTTCACTCCGGGACCTGGAAGATCTGGAATTTGATCTTTC 722
Db 2117 TACCATTAAGATATCTCAAGATTTCACTCCGGGACCTGGAAGATCTGGAATTTGATCTTTC 2176
QY 723 AGTAGTAAACCAACAGAGAGAGAGACCATTAAGGAAAGAAATTAAGTGACTTACA 782
Db 2177 AGTAGTAAACCAACAGAGAGAGAGACCATTAAGGAAAGAAATTAAGTGACTTACA 2236
QY 783 GGAGAGCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 842
Db 2237 GGAGAGCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2296
QY 843 TGATTAACCATTTGTCAAGAGAGTATACATGATTAATTAATTAATTAATTAATTAATTAATTA 902

Db 2297 TGATTAACCCATTGTCAAGAGATGATGATTAATTCAGC-ATGGGATTTCTCTAA 2355
QY ATAAAGCTTGGAAAGACCGGATTGACCAACTGTGAG-----AGCAATGACAGCTGCCGTCA 956
Db 2356 ATAAAGCTTGGAAAGACCGGATTGACCAACTTGGAGGCAATGACAGGCTGCCGTTCAGT 2415
QY 957 GTACAGATGACCTGCAGGCGGATATTTTCTGGGTAGATATTTGACGGTGTAGTTCGCTT 1016
Db 2416 ACCAGAGTGAAGCTGACAGGCGGTATTTGACTGGGTAGTATTTGACGGGTGTAGTTCGCTT 2475
QY 1017 CAATGCTCCAAATTTGGAAAGCATCTGAAACTGTCAAGACGACGATGATGTGCTAGAG 1076
Db 2476 CAATGCTCCAAATTTGGAAAGCATCTGAAACTGTCAAGACGACGATGATGAAAGCTTAAG 2534
QY 1077 CAATTTAAGCTGAGGCTTATCAACAGCAGATAG-AGTGTAGACGACTCATCAGCA 1135
Db 2535 CAATTTAAGCTGAGGCTTATCAACAGCAGATAGAAATGAAAGATGAAAGATCAATCAGCA 2594
QY 1136 GACGCTTTTGTCTAAGAAAGTAAACAGNAGAGATGACAAACAGACTGTTCTAGACCCAA 1195
Db 2595 GA-GCTTTTGTCTAAGAAAGTAAACAGNAGAGATGACAAACAGACTGTTCTAGACCCAT 2653
QY 1196 AATGAGATCTATATTGATAT-GGCTAGCCTGAGNAGAGATCATCAACAGACAGCATTA 1254
Db 2654 AATGAGATCTAATATTGATATGATGATAGCTGAGAGAGATCATCAACAGACAGCATTA 2713
QY 1255 ACTGAGAGGCTCTATTTAGCCNTGGGTGAGTTCACACATGCCCTGTATAGGCTCCG-- 1312
Db 2714 ACTGAGAGGCTCTATTTAGCCNTGGGTGAGTTCACACATGCCCTGTATAGGCTCCG 2773
QY 1313 CATGCTGACACACACGAGGCTTGTCTAATGAGCAGAAACCTGTTGAGAGACCCCTTA 1372
Db 2774 ATGGCTGACACACACGAGGCTTGTCTAATGAGCAGAAACCTGTTGAGAGACCCCTTA 2833
QY 1373 AGCATTGGAATTGAATCTTGCCAGACATCATGTGCTCCAAATGATATTTAGCCATCA 1432
Db 2834 AGCATTGGAATTGAATCTTGCCAGACATCATGTGCTCCAAATGATATTTAGCCATCA 2893
QY 1433 GTCACAGTGGAGCCGTTAATTAAGCAGAAATGATCTAATTTGAATCAAGTCAAGAGAG 1492
Db 2894 GTCACAGTGGAGCCGTTAATTAAGCAGAAATGATCTAATTTGAATCAAGTCAAGAGAG 2953
QY 1493 AGAAGCAAGCAACCTTCAGAAACAAGCTAGAGGTTTAAATCAACGCTGCCAAATGTTT 1552
Db 2954 AGAAGCAAGCAACCTTCAGAAACAAGCTAGAGGTTTAAATCAACGCTGCCAAATGTTT 3013
QY 1553 GGAATAAAACGAAACAAAGAAAGACAGAGCTGATGTGCTTGCCAGCCAAAGGCTT 1612
Db 3014 GGAATAAAACGAAACAAAGAAAGACAGAGCTGATGTGCTTGCCAGCCAAAGGCTT 3073
QY 1613 CCATGCGGAAATTTGAGGATTTGACAGAGTGGCTGACTGACACGAGAGGCTATGTTGGC 1672
Db 3074 CCATGCGGAAATTTGAGGATTTGACAGAGTGGCTGACTGACACGAGAGGCTATGTTGGC 3133
QY 1673 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGCAGCTTAATGCTCATATGGA 1732
Db 3134 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGCAGCTTAATGCTCATATGGA 3193
QY 1733 AGTCTGTGCTGCTTTGAGAGCTTAAGAAAGAAACATATAGAGTCTGATGAGAAAGGCCA 1792
Db 3194 AGTCTGTGCTGCTTTGAGAGCTTAAGAAAGAAACATATAGAGTCTGATGAGAAAGGCCA 3253
QY 1793 GCAGATGCTTGCAAGATGCCCCAAATCTGACAGAGCAAAATATTGACCAAGACAT 1846
Db 3254 GCAGATGCTTGCAAGATGCCCCAAATCTGACAGAGCAAAATATTGACCAAGACAT 3307

APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steilmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12468
LENGTH: 3987
TYPE: DNA
ORGANISM: Homo sapiens
US-10-198-846-12468

Query Match 39.7%; Score 807; DB 14; Length 3987;
Best Local Similarity 95.1%; Pred. No. 5,2e-200;
Matches 884; Conservative 0; Mismatches 40; Indels 6; Gaps 5;

QY 962 GATGACTGCAGCGGTATTTTCTGGGTAGATATTGACAGTGTAGTTCGCTTCAATG 1021
Db 46 GATGACTGCAGCGGTATTTGACTGGGTAGATATTGACAGTGTAGTTCGCTTCAATG 105
QY 1022 TCTCCAAATTTGGAACAGCATCTGAAAACCTGCAAGCAGCATTTGATGTCTAGACAA 1081
Db 106 TCTCCAAATTTGGAACAG-ATCTGAAAACCTGCAAGCAGCATTTGAAAGACTTAAGCAAT 164
QY 1082 TAAGTCTGAGGCTTATCAACAGACAGTAG-AGTGTAGACGACTCATCAGCAGACGC 1140
Db 165 TAAGTCTGAGGCTTATCAACAGCAGATAGAAATGAAAGATGATATCAACAGCAG-GC 223
QY 1141 TTTTGTAAAGAAAGTAAACAGNAGAGATGACAAACACACTGTTCTAGACCCAAATG 1200
Db 224 TTTTGTAAAGAAAGTAAACAGNAGAGATGACAAACACACTGTTCTAGACCCATTAATG 283
QY 1201 AACTGATATTGATAT-GGCTAGCCTTGAGNAGAGATCATCAACAGACGATTAACCTGG 1259
Db 284 AACTGATATTGATATGAGATAGCTTGAGAGAGATCATCAACAGACGATTAACCTGG 343
QY 1260 AGGCTCTCTATTAGCCNTGGGTGAGTTCACATGSCCTGTATGAGCTCTG--CATGC 1317
Db 344 AGGCTCTCTATTAGCCNTGGGTGAGTTCACATGSCCTGTATGAGCTCTGCGATGGC 403
QY 1318 TGAACACACCGAGGCTTGTCTAATGAGCAGAAACCTGTTGAGAGAGACCCCTAAAGCCA 1377
Db 404 TGAACACACCGAGGCTTGTCTAATGAGCAGAAACCTGTTGAGAGAGACCCCTAAAGCCA 463
QY 1378 TTGAATTTGAATCTTGCCAGACATCATGTGCTCCAAATGATATTTAGCCATCACTCA 1437
Db 464 TTGAATTTGAATCTTGCCAGACATCATGTGCTCCAAATGATATTTAGCCATCACTCA 523
QY 1438 CAGTGAAGCCGTTAATTAAGCAGAAATGATCTAATTTGAATCAAGTCAAGAGAGAGAG 1497
Db 524 CAGTGAAGCCGTTAATTAAGCAGAAATGATCTAATTTGAATCAAGTCAAGAGAGAGAG 583
QY 1498 CAAGCAACCTTCAGAACAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAA 1557
Db 584 CAAGCAACCTTCAGAACAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAA 643
QY 1558 AAAACAGAAAGAAAGAGAGCAGCTGATGTGCTTGCCAGGCAAAAGGTTTCATG 1617
Db 644 AAAACAGAAAGAAAGAGAGCAGCTGATGTGCTTGCCAGGCAAAAGGTTTCATG 703
QY 1618 GCGAAATTTGAGATTTTGACAGAGTGGCTGACAGACGAGAGGCTATCTGTTGGCATCTA 1677
Db 704 GCGAAATTTGAGATTTTGACAGAGTGGCTGACAGACGAGAGGCTATCTGTTGGCATCTA 763
QY 1678 AACCGCTGGAGGTTTAAACCGGAAACAGCCAAAGAGCAGCTTAATGCTCATATGAGATCT 1737

|||||
Db AACGCTGGAGGTTTACCGGAAACAGCAAGACAGCTTAAATGTCATATGAAAGTCT 823
Oy 1738 GTGCTGCTTTGAAAGCTTAAAGAAACATATAAGTCTGTGCAAGAAAGCCGACGA 1797
Db 824 GTGCTGCTTTGAAAGCTTAAAGAAACATATAAGTCTGTGCAAGAAAGCCGACGA 883
Oy 1798 TGCTTGAAGATGCCCAAAATCTGAGAGCAAAATATTGACCAAGACATTAATACTTGA 1857
Db 884 TGCTTGAAGATGCCCAAAATCTGAGAGCAAAATATTGACCAAGACATTAATACTTGA 943
Oy 1858 AAAAAAAAAAAAAAAAAAAAAAAAAAANAC 1887
Db 944 AAGAAAAATGGGAATCGGTGGAACCAAC 973

RESULT 6
US-10-198-846-12599/c
; Sequence 12599, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinhilber, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198, 846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12599
; LENGTH: 18529
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-12599

Query Match 32.1%; Score 652.2; DB 14; Length 18529;
Best Local Similarity 62.8%; Pred. No. 3.7e-159;
Matches 1132; Conservative 0; Mismatches 647; Indels 24; Gaps 7;

Oy 41 CCCCCCGAGCCCTTGAATATGATATCTTAAGGACGACAGAGAAACATCGCACTG 100
Db 6165 CCCTCTCCAGCCATGATCATGAGCAGCTCAGGACGAAACAAAGAGAAATGAGCAATTA 6106
Oy 101 CGTAGATGATAGCTGAACACAAAGCTCATATAGTAAATGAACAAACCTGGCC---A 156
Db 6105 AGGGAATATATGCTGAACACAAACCTTATATGCAAACTAATAAATAGGCCCAAA 6046
Oy 157 CAGTACTGAATGAGCCCTGGGAAAGGCTTTCTATCAAGAGAAATATGTCAGAGCCGAC 216
Db 6045 CTAAGAGAAATTAACCTGAGAGAGGAGAAATGTCGAAGAAATACCAAGAACAGAA 5986
Oy 217 ACCCTTAACATCACTTAAGAAAGATGTCAAAAAGCGTGTGTCGACCTGGATGAAGCC 276
Db 5985 AACATGATGCCCAAAATTAAGAGAGGAGTGCGCCAGCAGCCCTGGCTCTGGATGAAGCC 5926
Oy 277 ATTTCTCAATCAATC-----AGTTCCATGACAAATGATGATCAATCTTGAAGC 327
Db 5925 GTGTCCACGTCACACAGATTACAGAGTTTCAATGATTAATTAAGCCCTATGTTGAGCA 5866
Oy 328 CTGGAACGATCGTGAAGCTGTGAGCAGCAACCTCTATCTGTGAGAGGTTGAGAG 387
Db 5865 CTGGAAGATCTTCTCTCGCTGCTGTATGCCACACATGATCTCTGCTGAAGTGAAG 5806
Oy 388 ATCAAGAAACAGATGATGAATAATGAATGTGTCAATGACATGAAAGCTACAGCCG 447
Db 5805 ATCAAGAGATGATGATGATGAATAATGAATGTGTCAATGACATGAAAGCTACAGCCG 5746

Oy 448 TTGATGAACTCTTAAACAGAGGAGAGAGAAATGATGCTAGATCTGGGGAGCTGAT 507
Db 5745 TCCTTTGAGGCTTTGAAGCCGCTGTGAGAGAGAGCTTATGGAACATCTCAGAGAGCAGAC 5686
Oy 508 AAAGACATATCTGCAAGGCTGTTCAAGATTAAGCTTGAACAAATGTTTCAATTTGGAG 567
Db 5685 AAGATCTGCTGCAAAAGAAATCAGAGATTAATGATCAAAATGATCTTCTGCGGAG 5626
Oy 568 AACATACACACATCTGTGAGAGAGAGAGAGCCCAATCACTGATGTGATGAGTACAGCA 627
Db 5625 GACATCAAAAGCTCGGCTGAGAAACGAGAAATCAAAATTTCTTGATGTCTTGAATTTGCA 5566
Oy 628 GAAAGTTCTGTGTGATACATGATGATGATGATGATGATGATGATGATGATGATGATG 687
Db 5565 GAGAAATTTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5506
Oy 688 ATCCGGGACCTGGAAGATCTGGAATGATCTTCAATGATGATGATGATGATGATGATG 747
Db 5505 GTCCATGATCTGGAAGAGCCGAGGATGATCTTCCATCAACCAACAGAGTTGAAGCT 5446
Oy 748 GCAGAGACATTAAGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 807
Db 5445 GCTAGACTATTAAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 5386
Oy 808 CTAGATCTGAACTCATTCGCGCATGATGAGGAGGCTGATTAACCATTTGCAAGAGAT 867
Db 5385 CTGAGACAGATTTGATTTTTCCTGTGAGAGAACTGAGAGCTGAAATGAGAGAGAGC 5326
Oy 868 ATACATGATTAATTAATTCAGAGATGAGATTCCTTAATTAAGCTTGAAGAGCCGATG 927
Db 5325 ATGATGATGATGATTAATGCTGAGAGAACTGAGAGAGCTGAAATGAGAGAGAGC 5267
Oy 928 CCAACTGAG-----AGCAATGACAGCTGCTGATGATGATGATGATGATGATGATG 981
Db 5366 AAAAATTGAGAGATCTATGCAAGCTGCTGATGATGATGATGATGATGATGATGATG 5207
Oy 982 TTTCGGGTAGATTAATGAGAGTGTGTAAGTTGCTTCAATGATGATGATGATGATGATG 1041
Db 5206 TGATCTGATGATTAACATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 5148
Oy 1042 TCGAACTGTCAGACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1101
Db 5147 TCAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5088
Oy 1102 AGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1161
Db 5087 AGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5028
Oy 1162 NAGAGATGACAAACACATGTTCTAGCCCAAAATTAATTAATTAATTAATTAATTAAT 1220
Db 5027 ATGAGAGGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4968
Oy 1221 GCCTGAGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1280
Db 4967 ACCGTGGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4908
Oy 1281 GTCAATTCACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1338
Db 4907 GTCAATTCACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4848
Oy 1339 TAAATGAGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1398
Db 4847 TAAATGAGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4788
Oy 1399 ATCATGCTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1458
Db 4787 ACATGCTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4728
Oy 1459 CAGAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1518
Db 4727 CTGCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4668
Oy 1519 TAGAGGTTTAAATCAACGCTGCAAAATGTTTGGAAAAAACAAGAAAGAGCAGC 1578

Db 4667 TGGAAACATGAAACCAATGCTGGAGTCAAGTCTTACGAAACAGAGAGAGAGAGC 4608
Qy 1579 AGCTGATGATGCTGCTGGCCGACGACAAAGGTTCCATGGCGAAATTGAGATTGCGAC 1638
Db 4607 AGCTTCAGTCACTCTGACAGCAGGCGCTTCCAGATGAATTTGAGATTCTCT 4548
Qy 1639 AGTGGCTGACTGACAGGAGCGTCACTGTTGGCATTTAAACGCTGGAGGTTTACCG 1698
Db 4547 TGGAACTTAATGAAATGAGAGCCAGCTTCTGCACTTAAGCCCAAGAGAGACTTCT 4488
Qy 1699 AAACACCCAGAGAGAGCTTAATGCTCAATATGAAAGTCTGCTGCTTTGAAGCTTAAG 1758
Db 4487 AAACCTCTAGGAAACAGCTTGATACATATGAACTTAATCCAGCTGAAGCCAGG 4428
Qy 1759 AAGAAACATTAAGAGTCTGATGAGAAAGCCAGAGATGCTTGAAGATGCCAAAT 1818
Db 4427 AAGAGACTTAATCAACTTGAAGAGGAGAGACTCAATGCTTCTAAGCCGTGAGACT 4368
Qy 1819 CTG 1821
Db 4367 CTG 4365

RESULT 7
US-09-823-245A-445
/ Sequence 445, Application US/09823245A
/ Publication No. US20020039760A1
/ GENERAL INFORMATION:
/ APPLICANT: Wong, Gordon G.
/ APPLICANT: Clark, Hilary
/ APPLICANT: Fechtel, Kim
/ APPLICANT: Agostino, Michael J.
/ APPLICANT: Howes, Steven H.
/ APPLICANT: Resnick, Richard J.
/ APPLICANT: Gulukota, Kamalakar
/ APPLICANT: Graham, James R.
/ APPLICANT: Genetics Institute, Inc.
/ TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
/ FILE REFERENCE: GIN 6401
/ CURRENT APPLICATION NUMBER: US/09/823,245A
/ CURRENT FILING DATE: 2001-03-29
/ PRIOR APPLICATION NUMBER: 60/194,941
/ PRIOR FILING DATE: 2000-04-06
/ NUMBER OF SEQ ID NOS: 631
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 445
/ LENGTH: 2181
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-823-245A-445

Query Match 24.7%; Score 501; DB 9; Length 2181;
Best Local Similarity 62.0%; Pred. No. 4,4e-120;
Matches 895; Conservative 0; Mismatches 529; Indels 19; Gaps 6;

Qy 296 TCCATGACAGATAGATCAGATCCTTGAGAGCTGGAACGATCGTGAAGCTTGAAGC 355
Db 1 TCCATGATAAATTGAGCCATGTTGAGAGCACTGAGAACTTTCTCTCCGCTGCGTA 60
Qy 356 AGCCACCCCTTATCTCTGACAGAGTTGAGAAATCAAGAAACAGATCAGTGAATAAGA 415
Db 61 TGCACACACTGATCCCTGCTGAGAGTAGACAAAGATCAGAGAGTCAATCAGTGAATAAGA 120
Qy 416 ATGTGTCATGATGATGAGAAAGTACAGCCGTTGATGAACTTTAAACAGAGGAG 475
Db 121 GTCCACACCGTGAAGCTAGAAACCTGACGCCATCTTTGAGGCTTGAAGCGCCGTGAG 180
Qy 476 AGGAAATGATGCTAGATCTGGGGGAGCTGATAAGACATATCTGCCAAGCTGTTCAAG 535
Db 181 AGGAGCTTATTTGAGCAATCTCAGGAGACAGAAAGATCTGCTGCAAAAAGAAATCCAGG 240
Qy 536 ATAGCTTGACCAATGCTTTTCAATTTGGAGAACATTAACACACTGCTGGAGAGAGAGG 595

Db 241 ATAAATGATCAAAATGATATCTCTGGAGAGACATCAAGCTCGGCGTGAAGACGAG 300
Qy 596 AAGCCAACTACTGATGATGAGAGCTAGCAGAAAGTTGCTGCTGATCAATCAT 655
Db 301 AAATCAAAATTTCTTGATGCTGTTGAATTTGACAGAAATTTCTGATGACATGGCAGCTC 360
Qy 656 TGATAGTTACATTAAGATTAATCAAGATTTATCCGGAGCTGGAAGATCTCGAATTG 715
Db 361 TCTGACCAACATCAAAAGACCCAGGATATGTCATGATCTTGAAGAGCCAGGACTTGG 420
Qy 716 ATCTTCAAGTATTAACCAAGCAAGAAAGCAGCAGAGACATTAAGGAGAAATGATG 775
Db 421 ATCTTCCATCAATCAAAACAAAGGTTGAAGCTGCTGAGACTATTAAGAGAGACGATG 480
Qy 776 GACTACAGAGAGAGCTGATATAGTTATTAATTAACCTAGTTCTGAACATTTGGGAGCTG 835
Db 481 GTCTGATAGAGAGCTGAGGATTAATGCGATCTTTGAGACAGATTTGATTTTGGCTGTG 540
Qy 836 GAGAGCCTGATTAACCCATTTGTCAGAGAGATACATGATTAATTAATGAGAGATG 895
Db 541 GAGAACTGAGAAAGCTGAAATGAATTAATGCTTGGAGAACTTAAACGAAACATGG--- 597
Qy 896 TCTTAATTAAGCTTGGAAAGACCGGATTAACCACTTGAAGACATGAGCTGCCGTC 955
Db 598 ---AAGATAGGCTAGCAAAAATTG-----AAGATGCTATGACAGGCTGCTGACG 646
Qy 956 AGTACAGATGAGCTGAGAGCGGATTTTCTGGGTGATATGCAAGTGTGAAGTTCGCT 1015
Db 647 TATCAGAGAGCTTTCAAGCTATGTTGACTGACAGAAATTAACCTGATTAACCTTTC 706
Qy 1016 TCAATGCTCAATTTGAAACAGCATCTCGAACTGTCAAGCAGCAGATTAATGCTGAG 1075
Db 707 ACCAAGCCCCCTGTTGGCTGTG-ACCTGATCTGTTAAAGATCACTTAATGAATGA 765
Qy 1076 GCAATTTAAGTCTGAGGCTATCAACAGCAGATAGTGTGATGACTACTACGCA 1135
Db 766 GGAGTTCAAGTAAAGATGATACAGCAGCAAAATGAGAGAGAACCTTAATCACCAGG 825
Qy 1136 GACGCTTTGCTAAAGAAAGTAAAGAGAGTGAACAGCAACACGTTCTGAACCCAA 1195
Db 826 TGAATGATGTTAAAGAAAGCTATGATGAGACGACAGACATTAATGAGAACTACT 885
Qy 1196 AATGAGCTGATATGATATGCT-AGCTGAGAGAGAGATCATCAAGACAGCATTA 1254
Db 886 GACAGAACTCAAAACCTCTGGGAGAACCTGGGTGAGAAATTTGCCACGACAGCA 945
Qy 1255 ACTGAGGCTGCTATTAAGCCATGCTGATGCTTCAACATGCCCTGATGAGCTCTGCA 1314
Db 946 ACTAGAAAGGGCTCTGTTGGCCCTGCTGATGCTCAGATGCTTGAAGAACTAATGAG 1005
Qy 1315 T--GCTGACACACACGAGGGCTGCTAAGTGAAGAAACGTTGAGAGAGACCTTA 1372
Db 1006 TTGGCTGACTCATACCGAAAGAGTTGATGATCTCAAGAACATTAAGTGAAGACCCAA 1065
Qy 1373 AGCCATTTGAATTTGAATCTGCAAGCATGATGCTCCAAATGATGATTAAGCCCATCA 1432
Db 1066 AGTCATTTAAGTTGAGCTGCAAGACCATGCTTAATAAATGATGTTTGGCTCATCA 1125
Qy 1433 GTCCACAGTGAAGCCGTTAATTAAGCAGAAATGATCTTAATGAATCAAGTGCAGAGA 1492
Db 1126 AGCCACAGTGAAGAACGTCACAAAGCTGGCATGAGCTTCTTGAATCCAGTGTGAGAA 1185
Qy 1493 AGAAGCAACCACTTCAAGACAGTGAAGTTTAATCAAGCTGCAAAATGTTTT 1552
Db 1186 TGATGCAACAGCTTAAGAGCCGTTTGAAGAACATGAACCAATGCTGAGAGCTGATGTT 1245
Qy 1553 GGAAGAAACAGAAACAGAGACAGCAGCTGATGCTGCTGCGCAGACCAAGGCTT 1612
Db 1246 ACAGAAACAGAGAGAGAGAGAGCAGCTTCACTCACTCTGACAGACAGCCAGGCTT 1305
Qy 1613 CCATGGCGAAATTTGAGATTTGACAGAGTGGCTGACTGACAGAGAGCTGATCTGTTGGC 1672

Db 1306 CCACAGTGAATGAGA-TCCTCTTGGAATTACTAGATGAGAGCAGCTTTCTGC 1364
Qy 1673 ATCTAAACCGCTGGAGGTTTACCGGAACCCAGAGAGAGCTTAATGTCATATGGA 1732
Db 1365 ATCTAAACCCACAGAGAGACTTCTGAAACGTGAGGAGACGTTGATACATATAGT 1424
Qy 1733 AGT 1735
Db 1425 AAT 1427

RESULT 8
US-10-158-034-52
; Sequence 52, Application US/10158034
; Publication No. US20030219758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT210C1
; CURRENT APPLICATION NUMBER: US/10/158,034
; PRIORITY FILING DATE: 2002-05-31
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 2113
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-158-034-52

Query Match 24.5%; Score 497.8; DB 15; Length 2113;
Best Local Similarity 63.7%; Pred. No. 3e-119;
Matches 836; Conservative 0; Mismatches 466; Indels 11; Gaps 5;

Qy 518 CTGCCAAAGCTGTTCAAGATAGCTGACCAATGCTTTTCATTTGGAGAACTACACA 577
Db 62 CGGCAAAAGAAATCAGATTAATTTGATCAATGTTCTTCTGGAGAGACATCAAG 121
Qy 578 CACTGCTGGAAGAGAGGAGCCAACTACTGATGATGATGAGCTAGCAGAAAATTCT 637
Db 122 CTCGGGCTGAAGAGAGAAATCAATTTCTGATGCTCTTGAATAGCAGAGAGTTCT 181
Qy 638 GGTTGATCAATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 697
Db 182 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241
Qy 698 TGAAGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 757
Db 242 TGAAGAGCCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301
Qy 758 TAAGGAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 817
Db 302 TTAAGGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
Qy 818 AACCATTTGGCGGAGTGGGAGGCTGATAAACCATTGTAAGAGATGATGATGATGAT 877
Db 362 ATTGATTTTGGCTGAGAGAACTGGAAGCCGTAAGTGAAGAGATGATGATGATGAT 421
Qy 878 TAAATTCAGCAGTGGATTTCTTAATTAAGCTTGAAGAGCCGATGACCACTTGAG 937
Db 422 TGAATTAATGC-TTGGAGAACTTAACAAACATGGAAGAGAGGCTTGAAGAACTTGAG 480
Qy 938 -----AGCAATGACGCTGCGCTCAGTACAGATGAGCTGACGGCGGATTTTCTGGGTA 991
Db 481 GATGCTATGCAAGCTGCTGTCAGATCAGACACTCTTCAAGGCTATGTTGACTGAGCTA 540
Qy 992 GATATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1051
Db 541 GATTAACATGATGATTAATTAATCTTGCAACATGCCCCCTGTTGACATG-ACCTTAATACGT 599
Qy 1052 CAACAGCAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1111
Db 600 TAAAGATCAGTTAATGAATGAAGGAGCTTCAAGATGAAATTTTCAACAGCAAAATTGA 659

Qy 1112 GTGCTAGACGACCTCATCAAGAGAGCTTTTGTCTAAAGAAAGTAAACAGNAGAGTGA 1171
Db 660 GATGAGAGAGCTTAATATCAACAGAGGTGAATGATGTTAAAGAACTCTGATGAGAGGA 719
Qy 1172 CAACACACTGTTCTAGAACCCAAATAAGANCTGATATTGATAT-GGCTAGCCTGAGANA 1230
Db 720 CAGAGACTTATACAGAAACCACTGACAGAACTCAACCTCTGGAGAAACCTGGGGTGA 779
Qy 1231 GAGATCATCAACAGACATTAATCTGAGAGGTGCTCTATTAGCCNTGGTCTGTTCCA 1290
Db 780 GAAATATGCCACAGACAGCAAACTAAGAGGGGCTCTGTGGCCCTTGCTGATGTTCCA 839
Qy 1291 ACATGCCCTGATAGAGCTCTGCAT--GCTGACACACAGGAGGCTGTGCTAGTGAAGA 1348
Db 840 GCATGCTTTAGAGAACTTAATGAGTGGCTGACTCATACCGAAGATTTGATAGTCTCA 899
Qy 1349 GAAACTGTTGAGAGAGACCTTAAGCCATTGAATTTGAATTGCACTTGCAGCATCTGCT 1408
Db 900 GAGACCAATAGTGAAGACCCAAAGCTATTGAAGTTGAGCTGCAAAAGCACATGCTCT 959
Qy 1409 CCAAAATGATGATTTAGCCCATCATGTCACAGTGAAGCCGTTAATTAAGCAGAAATGA 1468
Db 960 AAAAATGATGTTTGGCTCATCAAGCCACAGTGAACAGTCAACAAAGCTGGCAATGA 1019
Qy 1469 TCTAATTTGAATCAAGTCAAGAGAGAGAGAGCAAGCACTTCAGAACAGCTAGAGTTT 1528
Db 1020 GCTTCTTGAATCCAGTCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 1079
Qy 1529 AATCAACGCTGCGCAAAATGTTTGGAAAAACAGAACAAAGAGAGAGAGAGAGAGAG 1588
Db 1080 GAACCAATGCTGGAGAGTCAATGTTTACAGAAACAGAGAGAGAGAGAGAGAGAGAG 1139
Qy 1589 TGCTTCGCGCAGGCGCAAGAGGTTTCATGCGCAATTTAGAGATTTGACAGAGTGC 1648
Db 1140 AACTCTGACAGAGGCCAGAGGCTTCCACAGTGAATTTGAAGATTTCTTGTGAATCTAC 1199
Qy 1649 TGACACGAGAGCTGATGTTGTTGATGATGATGATGATGATGATGATGATGATGAT 1708
Db 1200 TAGAATGAGAGAGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1259
Qy 1709 GAGACAGCTTAATGTCATATGAGAGTGTGCTGCTTGAAGACTTAAGAGAAACATA 1768
Db 1260 GGAACAGCTTGTATACATATGAACTCTTATTTCCACTGAAAGCAAGAGAGACTTA 1319
Qy 1769 TAAGAGTGTATGAGAGAGAGGCGCAGATGCTTTCAGATGCCCAAAATCTG 1821
Db 1320 TAATCAACTACTTGACAAAGGCGAGACTGATGCTTCTTAAGCCGTGACGACTCTG 1372

RESULT 9
US-10-158-034-32
; Sequence 32, Application US/10158034
; Publication No. US20030219758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT210C1
; CURRENT APPLICATION NUMBER: US/10/158,034
; PRIORITY FILING DATE: 2002-05-31
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 5453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-158-034-32

Query Match 24.5%; Score 497.4; DB 15; Length 5453;
Best Local Similarity 63.7%; Pred. No. 6.2e-119;
Matches 835; Conservative 0; Mismatches 465; Indels 11; Gaps 5;

QY 520 GCCAAGCTGTTCAGATTAAGCTTGACCAATGGTTTCTATTGGGAGAACATACACACA 579
 Db 1 GCAAAAGAAATCCAGAGTAATATGATCAATGGTATTTCTTGGGAGCATCAAGCT 60
 QY 580 CTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 639
 Db 61 CGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 640 TGTGATCAACATGTCATGTAGTATTAACCAATTAAGATCTCAAGATTTTCATCCGGAGCTG 699
 Db 121 TATGACATGGCAGCTCTCTGACACACATCAAGAGACCCAGATATTTGTCATGACTTG 180
 QY 700 GAAGATCTGGAGATTTGATCTTCAATGATTAACCAACAGAGAGAGAGAGAGAGAGAG 759
 Db 181 GAAGAGCCAGGAGATTTGATCTTCAATGATTAACCAACAGAGATTTGATGAGTCTGAGACTAT 240
 QY 760 AGGAGAGAAATTAATGAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819
 Db 241 AAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 QY 820 CTGATTCGGGATGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879
 Db 301 TTGATTTTTCCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 QY 880 AATTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 937
 Db 361 AATTAATGCTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
 QY 938 ----AGCAATGACGCTGCGCTGAGTACAGATGAGAGAGAGAGAGAGAGAGAGAGAG 993
 Db 420 TGTATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
 QY 994 TATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1053
 Db 480 TATACATGATTAATTAATCTGACACATGCCCCCTGTTGGAGCTG-ACCTTAATACGTTTA 538
 QY 1054 AGCAGAGATTTGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1113
 Db 539 AAGATGATTAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 598
 QY 1114 GGTACGATGATCTATCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1173
 Db 599 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658
 QY 1174 AACACATGTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1232
 Db 659 GAGACATTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718
 QY 1233 GAATCATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1292
 Db 719 AAATTTCCCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778
 QY 1293 ATGCCCTGATAGAGCTCTGAGAT--GCTGACACACAGAGAGAGAGAGAGAGAGAGAG 1350
 Db 779 ATGCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 838
 QY 1351 AACCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1410
 Db 839 GACCAATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 898
 QY 1411 AAAATGATGATTAAGCCATCAGTCAACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1470
 Db 899 AAAATGATGATTTGGCTCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 958
 QY 1471 TAATTAATCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1530
 Db 959 TTCTTAATCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1018
 QY 1531 ATCAACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1590
 Db 1019 ACCAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1078
 QY 1591 CCTTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1650

Db 1079 CTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1138
 QY 1651 ACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1710
 Db 1139 GAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1198
 QY 1711 AGAGCTTAATGATTCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1770
 Db 1199 AACAGCTGATACAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1258
 QY 1771 AGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1821
 Db 1259 ATCAACTACTTGAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1309

RESULT 10

US-10-198-846-4738
 ; Sequence 4738, Application US/10198846
 ; Publication No. US2003009974A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lillie, James
 ; APPLICANT: Xu, Yongyao
 ; APPLICANT: Wang, Youzhen
 ; APPLICANT: Steilmann, Kathleen
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; TITLE OF INVENTION: THERAPY OF BREAST CANCER
 ; FILE REFERENCE: MRI-049
 ; CURRENT APPLICATION NUMBER: US/10/198,846
 ; CURRENT FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/306,220
 ; NUMBER OF SEQ ID NOS: 14064
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4738
 ; LENGTH: 599
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 599
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-198-846-4738

Query Match 21.2%; Score 430.8; DB 14; Length 599;

Best Local Similarity 94.0%; Pred. No. 5.3e-102;

Matches 500; Conservative 0; Mismatches 26; Indels 6; Gaps 5;

QY 962 GATGAGCTCAGAGCGGATTTTCTGAGATATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
 Db 61 GATGAGCTCAGAGCGGATTTTGAATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 1022 TCTCAATTGGAACAGCATCTGGAACCTGCAAGCAGAGATGATGATGATGATGATGAT 1081
 Db 121 TCTCAATTGGAACAG-ATCTGGAACCTGCAAGCAGAGATGGAAGAGCTTAAGCAATT 179
 QY 1082 TAAAGTGAAGGCTATCAACAGAGATAG-AGTGTGAAGATGATGATGATGATGATGAT 1140
 Db 180 TAAAGTGAAGGCTATCAACAGAGATAGAAATGGAAGAGATGAAATCAATCAAGAGAG- 238
 QY 1141 TTTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 Db 239 TTTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 298
 QY 1201 AACTGATATGATAT-AGCTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1259
 Db 299 AACTGATATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358
 QY 1260 AGGAGCTATATTAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
 Db 359 AGGAGCTCTATTAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 418

QY	1318	TGACACACACCGAGGGCTTGGCTAGTAGACGAAACCTGTGGAGAGAACCCCTAAAGCA	1377
Db	419	TGACACACACCGAGGGCTTGGCTAGTAGACGAAACCTGTGGAGAGAACCCCTAAAGCA	478
QY	1378	TTGAAATTGAACCTGCCAAGCATCATGNCGCCAAAATGATGATTTACCCATCAATCCA	1437
Db	479	TTGAAATTGAACCTGCCAAGCATCATGNCGCCAAAATGATGATTTACCCATCAATCCA	538
QY	1438	CAGTGAAGCCGCTTAATTAAGCAGGAATGATCTTAATTGAATCAAGTCAGG	1489
Db	539	CAGTGAAGCCGCTTAATTAAGCAGGAATGATCTTAATTGAATCAAGTCAGG	590

RESULT 11
US-10-037-270-352
; Sequence 352, Application US/10037270
; Publication No. US20030104529A1

APPLICANT: Tang, Y. Tom
 APPLICANT: Liu, Chenghua
 APPLICANT: Aeuudi, Vinod
 APPLICANT: Zhang, Jie
 APPLICANT: Ren, Feiyan
 APPLICANT: Chen, Rui-hong
 APPLICANT: Zhao, Qing A.
 APPLICANT: Wehrman, Tom
 APPLICANT: Xue, Aйдong J.
 APPLICANT: Yang, Yonghong
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Zhou, Ping
 APPLICANT: Ma, Yundong
 APPLICANT: Wang, Dunhui
 APPLICANT: Wang, Zhiwei
 APPLICANT: Tillinghast, John
 APPLICANT: Dmanac, Radje T.
 TITLE OF INVENTION: No. US20030104529a1el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 784CIP2B
 CURRENT APPLICATION NUMBER: US/10/037,270
 CURRENT FILING DATE: 2002-01-04
 PRIOR APPLICATION NUMBER: 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/488,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ. ID NOS: 1104
 SOFTWARE: pc_FL_genes Version 1.0
 SEQ ID NO 352
 LENGTH: 4495
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (307)..(3006)
 US-10-037-270-352

	Query Match	Similarity	13.1%	Score	266.8	DB	15	Length	4495	
	Best Local	Similarity	93.9%	Pred.	No.	1,2e-58				
	Matches	277	Conservative	0	Mismatches	18	Indels	0	Gaps	0
QY	1593	TTGGCCACAGGCCAAGGTTCCATCGCGCAATTTGAGATTTCACACAGTGGCTGACTGAC	1655							
Db	172	TTACTTCAGGCCCAAGGGTTCCATCGCGCAAAATTTGAGATTTCACACAGTGGCTGACTGAC	231							
QY	1653	ACGAGCGCTCATCTGTTGSCATCTTAAACCGCTGGAGAGTTTACCGGAAACAGCCCAAGAG	1712							
Db	232	ACGAGCGCTCATCTGTTGSCATCTTAAACCGCTGGAGAGTTTACCGGAAACAGCCCAAGAG	291							
QY	1713	CAGCTTAATGTCATTAATGGAAGTCGTGCTGCTTGAAGCTTAAGAAAGAAACATTAAG	1772							
Db	292	CAGCTTAATGTCATTAATGGAAGTCGTGCTGCTTGAAGCTTAAGAAAGAAACATTAAG	351							
QY	1773	AGCTGATGACAAAGAGCCAGACGATGCTTGCAGATGCCCAAAATTCGACAGACAAT	1833							

Db 352 AGCTCATGCAAAAGGCCAGACAGTGCCTGCGAAATGCCAAATCTCGAGACCAAT 412

Qy 1833 ATTGACCAAGACATAATTACTTTGAAAAAATTTTTTTTTTTTTTTTTTTTNC 1887

Db 412 ATTGACCAAGACTTAATTACTTTGAAAAAATGCGAATGCGTGGAAACCAATC 466

RESULT 12
US-10-117-722-352
; Sequence 352, Application US/10117722
; Publication No. US20030219744A1

```

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Drmanac, Radje T.
TITLE OF INVENTION: No. US2003021974A1el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP28CIP
CURRENT APPLICATION NUMBER: US/10/117,722
CURRENT FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/468,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 352
LENGTH: 4495
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (307)..(3006)
US-10-117-722-352

```

```

Query March 13.14; Score 266.8; DB 15; Length 4495;
Beet Local Similarity 93.9%; Pred. No. 1.2e-58;
Matches 277; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Dy 1593 TTGGGCCAGGCCAAGGGTTTCATGCGCAATTGNGATTTGCGACGATGGCTGACTGAC 1652
Dy 172 TTACTTCAGGCCAAGGGTTTCATGCGGGAATTGAGATTTTGCGACGATGGCTGACTGAC 231
Dy 1653 ACGGAGCGTCATCTGTTGGCATCTTAAACCGCTGGGAGGTTTACCGGAAAACAGCCAAAGAG 1712
Dy 232 ACGGAGCGTCATCTGTTGGCATCTTAAACCGCTGGGAGGTTTACCGGAAAACAGCCAAAGAG 291
Dy 1713 CAGCTTAATGTCATATGCAATGCAAGTCTGTGCTCCCTTTGAAAGCTTAAAGAAAGAAACATATAG 1772
Dy 292 CAGCTTAATGTCATATGCAATGCAAGTCTGTGCTCCCTTTGAAAGCTTAAAGAAAGAAACATATAG 351
Dy 1773 AGCTGATGCGAAGAGCCACAGACATGCTTGCAAGATGCCCAAAATCTGCGAGAGCAAAAT 1832
Dy 352 AGCTGATGCGAAGAGCCACAGACATGCTTGCAAGATGCCCAAAATCTGCGAGAGCAAAAT 411
Dy 1833 ATTGACCAAGACATTAATTAACCTTGAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTT 1887
Dy 412 ATTGACCAAGACATTAATTAACCTTGAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTT 466

RESULT 13
US-09-864-408A-111
; Sequence 111, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: NO. US20040009474A1 Human Polynucleotides and Polypeptides Enc
; FILE REFERENCE: 21402-012

```

;; CURRENT APPLICATION NUMBER: US/09/864,408A
;; CURRENT FILING DATE: 2001-05-24
;; PRIOR APPLICATION NUMBER: 60/206,690
;; PRIOR FILING DATE: 2000-05-24
;; NUMBER OF SEQ ID NOS: 9068
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 111
;; LENGTH: 344
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-864-408A-111

Query Match 12.9%; Score 262.8; DB 11; Length 344;
Best Local Similarity 92.4%; Pred. No. 3.6e-58;
Matches 318; Conservative 0; Mismatches 21; Indels 5; Gaps 4;

QY 1039 ATCTCGAAATCTGTCAGACAGCAGATTTGATGCTAGAGCAATTTAACTCTGAGGCTATC 1098
DB 2 ATCTCGAAATCTGTCAGACAGCAGATTTGAAAGCTAAAGCTAAATTAAGTCTGAGGCTATC 61
QY 1099 AACAGCAGATAG-AGTGTGTCAGACTGATCATGACGACAGCGCTTTTGTCTAAAGAAAGTA 1157
DB 62 AACAGCAGATAGAAATGAAAGACTGAAATCATCAAGCAGA-GCTTTTGTCTAAAGAAAGTA 120
QY 1158 ACAGNAGAGAGTACAAACACACTGTTCTAGACCCAAAAATGAGCTGATTTGATAT-G 1216
DB 121 ACAGNAGAGAGTACAAACACACTGTTCTAGACCCAAATTAATGAACTGAAATTTGATATGG 180
QY 1217 GCTAGCTGAGNAGNAGAAATCATCAACAGCAGATTTAACTGAGGGTGTCTATTATGCC 1276
DB 181 GATAGCGCTGAGAGAGAAATCATCAACAGCAGATTTAACTGAGGGTGTCTATTATAGCC 240
QY 1277 NTGGGTCACTTCCAAACATGCGCTGTATGAGCTCTG--CATGCTGACACACACCGAGGGC 1334
DB 241 TTGGGTCACTTCCAAACATGCGCTGTATGAGCTCTGAGCTGTGACACACACCGAGGGC 300
QY 1335 TTGCTAAGTGAGCAGAAACCTGTTGAGAGAGACCTTAAAGCCAT 1378
DB 301 TTGCTAAGTGAGCAGAAACCTGTTGAGAGAGACCTTAAAGCCAT 344

RESULT 14
US-10-043-487-126
;; Sequence 126, Application US/10043487
;; Publication No. US2003055220A1
;; GENERAL INFORMATION:
;; APPLICANT: HYBRIGENICS
;; APPLICANT: PIERRE, LEBRAIN
;; TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptide
;; FILE REFERENCE: B4778A
;; CURRENT APPLICATION NUMBER: US/10/043,487
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/261,130
;; PRIOR FILING DATE: 2001-01-12
;; NUMBER OF SEQ ID NOS: 561
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 126
;; LENGTH: 1721
;; TYPE: DNA
;; ORGANISM: Shigella flexneri
US-10-043-487-126

Query Match 11.6%; Score 235.6; DB 14; Length 1721;
Best Local Similarity 66.6%; Pred. No. 1.1e-50;
Matches 337; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 1316 GCTGACACACACCGAGGCTTGTCTAAGTACAGAAAACCTGTTGAGAGAGACCTTAAGC 1375
DB 1 GCTGACACACACCGAGAGAGTGTCTAAGTACAGAAAACCTGTTGAGAGAGACCTTAAGT 60
QY 1376 CATTGAATTTGAATCTTGCACAGACATCATGTCTCCAAATGATGATATTAGCCCATGATC 1435

DB 61 CATTGAATTTGAATCTTGCACAGACCATCTCTAAAAATGATGTTTGGCTCATCAAGC 120
QY 1436 CACAGTGAAGCGCTTAATTAAGCAGGAATATCTAATTGAATCAAGTGCAGAGAGC 1495
DB 121 CACAGTGAAGCGCTTAATTAAGCAGGAATATCTAATTGAATCAAGTGCAGAGAGC 180
QY 1496 AGCAAGCAACTTTCAGAACCAAGCTAGAGGTTTAAATCAACGCTGCAAAATGTTTGGC 1555
DB 181 TGCAGCAGCTTAAGAGAGCGTTTGAACCATGAACCAATGCTGGAGATGATGTAC 240
QY 1556 AAAACAGAACCAAGAGACAGAGCTGATGCTGCTTCCCGCAGCCAAAGGCTTCCA 1615
DB 241 GAAAACAGAGAGAGAGAGCAGAGCTTCAATCAATCTGACAGCAGGCCAGGCTTCCA 300
QY 1616 TGGCGAAATTTGAGATTTGCAGAGTGGCTGACTGACAGAGAGCGCATCTGTTGGCATC 1675
DB 301 CAGTGAATTTGAAGATTTCTCTTGAACCTTACTGAAGAGAGAGCAGCTTCTCATC 360
QY 1676 TAAACCGCTGGAGGTTTACCGGAAACAGCAGAGAGCTTAAATGTCATATGAGAT 1735
DB 361 TAAAGCCACAGAGAGAGCTTCTGAAATCTGAGGAAACAGCTTGAATACATATGAGAT 420
QY 1736 CTGTGCTGCTTGAAGCTTAAAGAAAGAAACATTAAGTCTGATGACAGAAAGCCAGCA 1795
DB 421 CTATTCACAGCTGAAGCCAGAGAGAGACTTATATCAATTAATCAATTTGACAGAGGCGACT 480
QY 1796 GATGCTTGAAGATGCGCCAAATCTG 1821
DB 481 CATTCTTGAAGCTGAGCAGACTCTG 506

RESULT 15
US-10-094-749-272
;; Sequence 272, Application US/10094749
;; Publication No. US20030219741A1
;; GENERAL INFORMATION:
;; APPLICANT: ISOGAL, TAKAO
;; APPLICANT: SUGIYAMA, TOMOYASU
;; APPLICANT: OTSUKI, TETSUJI
;; APPLICANT: WAKAMATSU, AI
;; APPLICANT: SATO, HIROYUKI
;; APPLICANT: ISHII, SHIZUKO
;; APPLICANT: YAMAMOTO, JUN-ICHI
;; APPLICANT: ISONO, YUUKO
;; APPLICANT: HIO, YURI
;; APPLICANT: OTSUKA, KAORU
;; APPLICANT: NAGAI, KEIICHI
;; APPLICANT: IRIE, RYOTARO
;; APPLICANT: TAMECHIKI, ICHIRO
;; APPLICANT: SEKI, NAOHITO
;; APPLICANT: YOSHIKAWA, TSUTOMU
;; APPLICANT: OTSUKA, MOTOMYUKI
;; APPLICANT: NAGAHARI, KENJI
;; APPLICANT: MASUHO, YASUHIKO
;; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;; FILE REFERENCE: 08435/0160
;; CURRENT APPLICATION NUMBER: US/10/094,749
;; CURRENT FILING DATE: 2002-03-12
;; PRIOR APPLICATION NUMBER: 60/350,435
;; PRIOR FILING DATE: 2002-01-24
;; PRIOR APPLICATION NUMBER: JP 2001-328381
;; PRIOR FILING DATE: 2001-09-14
;; NUMBER OF SEQ ID NOS: 3381
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 272
;; LENGTH: 3684
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-094-749-272

Query Match 8.8%; Score 177.8; DB 15; Length 3684;
Best Local Similarity 92.5%; Pred. No. 2.1e-35;
Matches 210; Conservative 0; Mismatches 12; Indels 5; Gaps 2;

```

OY      8 GCTGACAGAAACACATCATCTCTCAGC-TCCGCGCCGAGCCCTTGAATATGATAC 66
Db      3458 GCTGACAGAAACACATCATCTCTCAGCCTTCCGCGCCGAGCCCTTGAATATGAAAC 3517
OY      67 TCTAAGGCAGCAGCAGAGAAACATCGCACTGCGTAGTTGATAGCTGAACACAAGCC 126
Db      3518 TCTAAGGCAGCAGCAGAGAAACATCGCACTGCGTAGTTGATAGCTGAACACAAGCC 3577
OY      127 TCATATAGATAGATGAACAAACTCG---CCAAGTACTGAATGAGCCCTGGGGAGG 182
Db      3578 TCATATAGATAGATGAACAAACTCGGCCACAGTTACTGGAATTGAGCCCTGGGGAGG 3637
OY      183 CTTTCTATCCAGAGAGATATGTGGCAGCGGACACCCCTTACAGTC 229
Db      3638 CTTTCTATCCAGAGAGATATGTGGCAGCGGACACCCCTTACAGTC 3684

```

Search completed: January 12, 2005, 15:13:14
 Job time : 1090 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2005, 07:18:34 : Search time 193 Seconds

(without alignments)
7479.856 Million cell updates/sec

Title: US-10-089-887-46

Perfect score: 2031

Sequence: 1 ttgcacatgcgcacagaaca.....tngagagctcccaagcgc 2031

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	266.8	13.1	4495	4 US-09-620-312D-352	Sequence 352, App
2	250.2	12.3	278	4 US-09-513-999C-36133	Sequence 36133, A
3	93.8	4.6	1024	4 US-09-338-475C-9	Sequence 9, Appl
4	69.4	3.4	1613	3 US-08-938-830-28	Sequence 28, Appl
5	67	3.3	7218	1 US-08-232-463-14	Sequence 14, Appl
6	54.8	2.7	2311	4 US-09-720-317A-19	Sequence 19, Appl
7	52.8	2.6	2519	4 US-09-380-287A-9	Sequence 9, Appl
8	50.4	2.5	3472	4 US-09-873-737A-5	Sequence 5, Appl
9	50.2	2.5	1024	4 US-09-338-475C-37	Sequence 37, Appl
10	50	2.5	2365	3 US-09-249-687A-5	Sequence 5, Appl
11	50	2.5	2365	3 US-09-249-687A-18	Sequence 18, Appl
12	50	2.5	2365	3 US-09-363-316B-5	Sequence 5, Appl
13	50	2.5	2365	3 US-09-363-316B-23	Sequence 23, Appl
14	48.6	2.4	1493	1 US-08-340-820-24	Sequence 24, Appl
15	48.6	2.4	1493	1 US-08-593-535-24	Sequence 24, Appl
16	48.6	2.4	1638	4 US-09-591-095-26	Sequence 26, Appl
17	48.6	2.4	8607	4 US-10-204-708-72	Sequence 72, Appl
18	46.8	2.3	1459	4 US-09-537-654-3	Sequence 3, Appl
19	46.8	2.3	1663	4 US-09-464-535-43	Sequence 43, Appl
20	46.8	2.3	2527	4 US-09-244-805-29	Sequence 29, Appl
21	46.4	2.3	830	1 US-08-688-609-1	Sequence 1, Appl
22	46.4	2.3	1024	3 US-09-002-832-1	Sequence 1, Appl
23	46.2	2.3	1024	4 US-09-328-475C-36	Sequence 36, Appl
24	46.2	2.3	1069	4 US-09-205-258-74	Sequence 74, Appl
25	46.2	2.3	2799	4 US-09-181-339-2	Sequence 2, Appl
26	46.2	2.3	3124	4 US-09-734-030-1	Sequence 1, Appl
27	46.2	2.3	3124	4 US-10-153-921-1	Sequence 1, Appl

28	45.8	2.3	997	4 US-09-907-794A-376	Sequence 376, App
29	45.8	2.3	997	4 US-09-905-125A-376	Sequence 376, App
30	45.8	2.3	997	4 US-09-902-775A-376	Sequence 376, App
31	45.8	2.3	997	4 US-09-906-700-376	Sequence 376, App
32	45.8	2.3	997	4 US-09-903-603A-376	Sequence 376, App
33	45.6	2.2	1361	4 US-09-489-847-64	Sequence 64, Appl
34	45.6	2.2	1674	4 US-10-140-002-453	Sequence 453, App
35	45.6	2.2	2394	4 US-09-800-729-13	Sequence 33, Appl
36	45.4	2.2	3396	3 US-08-989-299-1	Sequence 1, Appl
37	45.4	2.2	3396	4 US-10-158-847-141	Sequence 141, App
38	45.4	2.2	3396	4 US-09-407-427-1	Sequence 1, Appl
39	45.2	2.2	708	4 US-09-270-767-13081	Sequence 13081, A
40	45.2	2.2	1024	4 US-09-338-475C-35	Sequence 35, Appl
41	45.2	2.2	2447	2 US-09-014-969-14	Sequence 14, Appl
42	45	2.2	52	4 US-09-807-784B-8	Sequence 8, Appl
43	45	2.2	931	4 US-09-482-273-31	Sequence 31, Appl
44	45	2.2	1801	4 US-09-709-103-3	Sequence 3, Appl
45	45	2.2	1801	4 US-09-439-410A-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-620-312D-352
Sequence 352, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyen
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19, 317
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 352
LENGTH: 4495
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: (307)..(3006)
US-09-620-312D-352

Query Match 13.1% Score 266.8; DB 4; Length 4495;
Best Local Similarity 93.9%; Pred. No. 1.1e-64;
Matches 277; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1593 TTGGCCAGGCGCAAGGTTTCATGCGGAATTGAGATTTCAGCAGTGGCTGACTGAC 1652
DB 172 TTACTTCAGGCGCAAGGTTTCATGCGGAATTGAGATTTCAGCAGTGGCTGACTGAC 231


```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,830
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/798419
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1066P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1613 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
;
US-08-938-830-28

Query Match          3.4%; Score 69.4; DB 3; Length 1613;
Best Local Similarity 87.9%; Pred. No. 2.7e-09;
Matches 87; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 1896 GAATCTAGATTCAAGCGCGCGCTGATTTAGACCTGCGCGCGCGCGCTGAGCCC 1955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 GAAAAGAGATTCAGCGCGCGCTGATTTAGACCTGCGCGCGCGCGCGCGCGCG 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1956 TATAGTAGTGCTATTAGATGGAATCACTAGTGGCGCG 1994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 105 TATAGTAGTGCTATTAGATGGAATGCGAATTTGCGAG 67
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
```

```

; TELEEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpc-F18
;
US-08-232-463-14

Query Match          3.3%; Score 67; DB 1; Length 7218;
Best Local Similarity 7.1%; Pred. No. 3e-08;
Matches 31; Conservative 23; Mismatches 13; Indels 0; Gaps 0;

QY 1492 AAGAGCAAGCACTTCAAGCAAGCTTAAATCAAGCTGCGCAAAATGTT 1551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1457 AAGAGATGAGAAATTTGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1552 TCGAAAAACGAACAAGACAGACGCTGATGTCCTGCGCCAGCCAAAGGT 1611
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1397 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1612 TCATGCGGAATGAGATTTGACAGAGTGCTGACAGACGAGCTCATCTGTG 1671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1337 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1672 CATCTAAACCGCTGGAGGTTTACCGGAAACAGCAAGACAGACCTTAATGTCATATG 1731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1277 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1732 AAGTCGTGCTGCTTGAAGCTAAAGAAACATTAAGAGCTGATGCAAGAGCC 1791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1217 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1792 AGCAGATGCTTGACAGATGCGCAAAATCTGAGAGACAAATATGACAGACATAATA 1851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1157 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1852 ACTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1097 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1912 CGGCGCTGAATTTCT 1926
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1037 CTCGAATTAATTTCT 1023
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-09-720-317A-19
; Sequence 19, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 2311
; TYPE: DNA
; ORGANISM: Triticum aestivum
;
US-09-720-317A-19

Query Match          2.7%; Score 54.8; DB 4; Length 2311;
Best Local Similarity 58.3%; Pred. No. 4.3e-05;
Matches 95; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
```



```

: APPLICANT: Ford, John
: APPLICANT: Yeung, George
: TITLE OF INVENTION: EGF MOTIF PROTEIN MATERIALS AND METHODS
: FILE REFERENCE: 28110/35852
: CURRENT APPLICATION NUMBER: US/09/353,316B
: CURRENT FILING DATE: 1998-07-28
: PRIOR APPLICATION NUMBER: US 09/249,697
: PRIOR FILING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: US 08/968,800
: PRIOR FILING DATE: 1997-11-22
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 23
: LENGTH: 2365
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (205)..(1863)
: US-09-363-316B-23

```

```

Query Match      2.5%; Score 50; DB 3; Length 2365;
Best Local Similarity 98.4%; Pred. No. 0.00098;
Matches 61; Conservative 0; Mismatches 0; Indels 1; Gaps 1#
OY      1927 AGACCTGCGCGGCGCGCGCTCGAGCCCTATAGTACGTGATAGGATGGATCACTA 1986
          |||||
Db       61  AGACCTGCGCGGCGCGC-CGCTCGAGCCCTATAGTACGTGATAGGATGGATCACTA 3
OY      1987 GT 1988
          ||
Db       2  GT 1

```

RESULT 14
 US-08-340-820-24
 : Sequence 24. Application US/08340820
 : Patent No. 3512460
 : GENERAL INFORMATION:
 : APPLICANT: NARUO, Ken-ichi
 : APPLICANT: SEKO, Chisako
 : APPLICANT: KUROKAWA, Tautomu
 : APPLICANT: KONDO, Tatsuya
 : TITLE OF INVENTION: GLIA ACTIVATING FACTOR AND ITS
 : TITLE OF INVENTION: PRODUCTION
 : NUMBER OF SEQUENCES: 27
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
 : ADDRESSER: CUSHMAN
 : STREET: 130 Water Street
 : CITY: Boston
 : STATE: Massachusetts
 : COUNTRY: US
 : ZIP: 02109
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/340,820
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/07/835,713
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: CONLIN, David G.
 : REGISTRATION NUMBER: 27026
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (617)523-3400
 : TELEFAX: (617)523-6440
 : TELEX: 200291 STRE UR

```

? INFORMATION FOR SEQ ID NO: 24:
?
? SEQUENCE CHARACTERISTICS:
?     LENGTH: 1493 base pairs
?     TYPE: nucleic acid
?     STRANDEDNESS: double
?     TOPOLOGY: linear
?
? MOLECULE TYPE: cDNA to mRNA
?
? HYPOTHETICAL: NO
?
? ANTI-SENSE: NO
?
? ORIGINAL SOURCE:
?     ORGANISM: Homo sapiens
?     HAPLOTYPE: 2n
?     TISSUE TYPE: skin
?     CELL TYPE: fibroblast
?
? IMMEDIATE SOURCE:
?     LIBRARY: Human foreskin cDNA library
?
? CLONE: pGAF1
?
? JS-08-340-820-24

```

	Query Match	Similarity	Score	DB 1	Length
	Best Local	58.3%	Pred. No. 0.0019		1493
	Matches	84	Conservative	0	Mismatches 60; Indels 0; Gaps 0;
QY	1745	CTTTGAAGCTAAAGAGAAACATTAAGAAGTCTGATGCAGAAAGCCAGCAGATGCTTGC	1804		
Db	1349	CTTAGTAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAGTTAAATTATTATTAAG	1408		
QY	1805	AAGATGCCCAAAATCTGCAGAGACAAATATTGACCAAGACATAAATACTTGAAGAAAAA	1866		
Db	1409	AAATTCCAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1468		
QY	1865	AAAAAAAAAAAAAAAAAAAAACA	1888		
Db	1469	AAAAAAAAAAAAAAAAAAAAA	1492		

RESULT 15
 US-08-593-535-24
 Sequence 24, Application US/08593535
 Patent No. 5622928
 GENERAL INFORMATION:
 APPLICANT: NARUO, Ken-ichi
 APPLICANT: SEKO, Chiako
 APPLICANT: KUROKAWA, Tsutomu
 APPLICANT: KONDO, Tatsuya
 TITLE OF INVENTION: GLIA ACTIVATING FACTOR AND ITS
 TITLE OF INVENTION: PRODUCTION
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
 ADDRESSEE: CUSHMAN
 STREET: 130 Water Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: US
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/593,535
 FILING DATE: 24-JAN-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/835,713
 FILING DATE: 12-FEB-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: CONLIN, David G.
 REGISTRATION NUMBER: 27026
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)523-3400


```

1 TELEFAX: (617)523-6640
2
3 TELE: 200291 STRE UR
4
5 INFORMATION FOR SEQ ID NO: 24:
6
7 SEQUENCE CHARACTERISTICS:
8
9 LENGTH: 1493 base pairs
10
11 TYPE: nucleic acid
12
13 STRANDEDNESS: double
14
15 TOPOLOGY: linear
16
17 MOLECULE TYPE: cDNA to mRNA
18
19 HYPOTHEITICAL: NO
20
21 ANTI-SENSE: NO
22
23 ORIGINAL SOURCE:
24
25 ORGANISM: Homo sapiens
26
27 HAPLOTYPE: 2n
28
29 TISSUE TYPE: skin
30
31 CELL TYPE: fibroblast
32
33 IMMEDIATE SOURCE:
34
35 LIBRARY: Human foreskin cDNA library
36
37 CLONE: pGAF1
38
39 JS-08-593-535-24

```

Query Match	2.4%	Score 48.6	DB 1	Length 1493
Best Local Similarity	58.3%	Pred. No. 0.0019		
Matches 84	Conservative 0	Mismatches 60	Indels 0	Gaps 0

QY 1745 TTTTGAAGCTAAAGAGAAACATTTAAAGCTATGCGAAGGCGCAGAGATGCTGC 1804

Db 1349 CTTGTGTAAAAAATATAAAAAATATAAAAAATATAAAAAATTAATTTATTAAG 1408

QY 1805 AAGATGCCCAAAATCTGCAGAGACAAATATTGACCAACACATAAATCTTGAAAAAA 1864

Db 1409 AAAATTCGCAAAAAAATATAAAAAATATAAAAAATATAAAAAATATAAAAAA 1468

QY 1865 AAAAAAAAAAAAAAAAAAAACA 1888

Db 1469 AAAAAAAAAAAAAAAAAAAAAA 1492

Search completed: January 12, 2005, 12:52:48
Job time : 195 secs

This Page Blank (uspto)